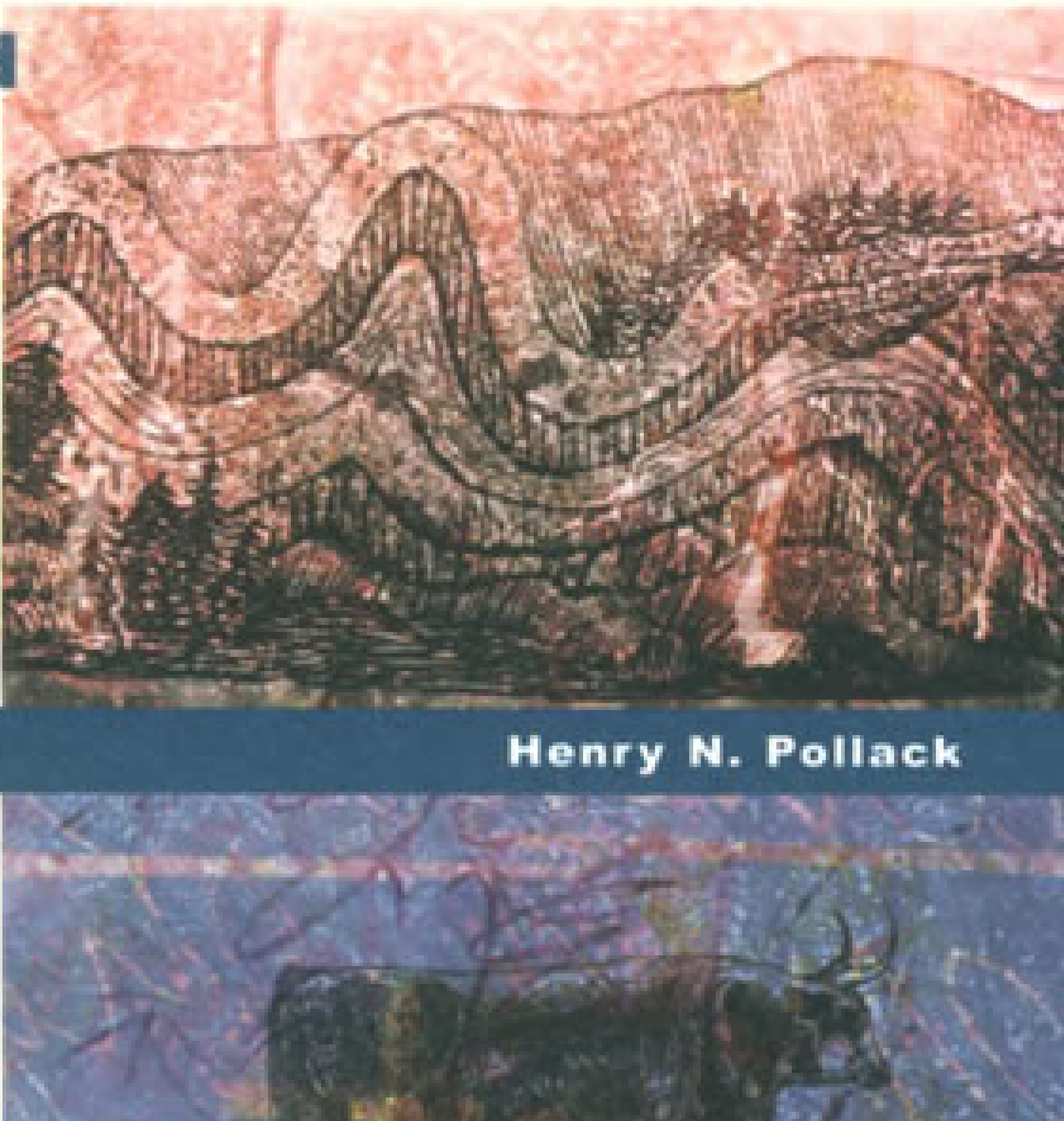
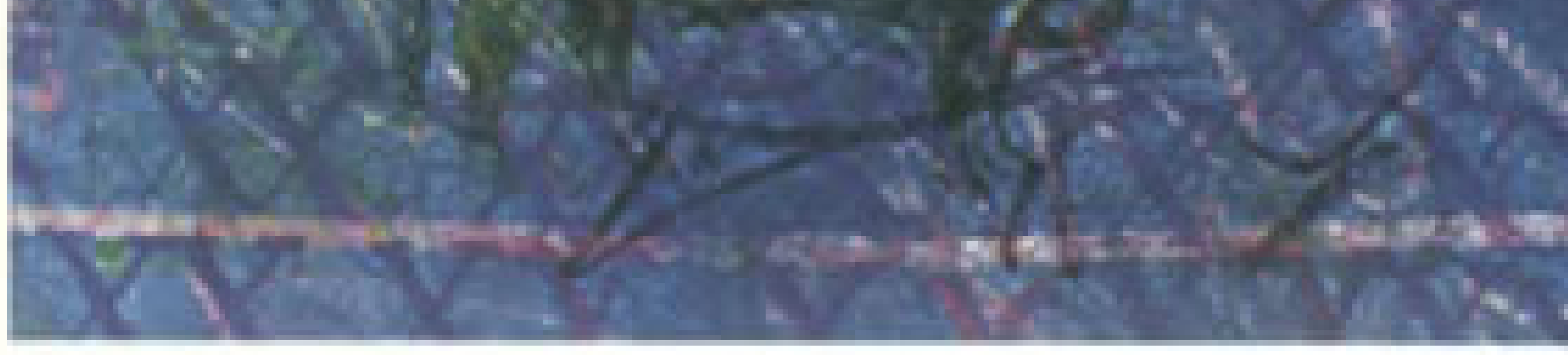
开放人文

不确定的科学与不确定的世界

Uncertain Science... Uncertain World



[美] 亨利·N·波拉克 著　李萍萍 译　Henry N. Pollack



上海世纪出版集团

不确定的科学与不确定的世界

[美] 亨利 · N · 波拉克 著　　李萍萍 译

读书中文
精品奉献
http://rbook.net/bbs/

世纪出版集团　上海科技教育出版社

图书在版编目(CIP)数据

不确定的科学与不确定的世界／（美）波拉克（Pollack，H.N.）著；李萍萍译.—上海：上海科技教育出版社，2005.9

（世纪人文系列丛书）

ISBN 7－5428－3796－6

Ⅰ.不… Ⅱ.①波… ②李… Ⅲ.不确定系统－研究 Ⅳ.N94

中国版本图书馆 CIP 数据核字（2005）第 099072 号

责任编辑 潘 涛 傅 勇

装帧设计 陆智昌

不确定的科学与不确定的世界

［美］亨利·N·波拉克 著

李萍萍 译

出　　版 世纪出版集团 上海科技教育出版社

（200235 上海冠生园路 393 号 www.ewen.cc）

发　　行 上海世纪出版集团发行中心

印　　刷 商务印书馆上海印刷股份有限公司

开　　本 635×965 mm 1／16

印　　张 16

插　　页 4

字　　数 195 000

版　　次 2005 年 9 月第 1 版

印　　次 2005 年 9 月第 1 次印刷

ISBN 7－5428－3796－6／N·664

图　　字 09－2004－018 号

定　　价 22.00 元

世纪人文系列丛书编委会

主任

陈　昕

委员

丁荣生　王一方　王为松　王兴康　包南麟　叶　路
张晓敏　张跃进　李伟国　李远涛　李梦生　陈　和
陈　昕　郁椿德　金良年　施宏俊　胡大卫　赵月瑟
赵昌平　翁经义　郭志坤　曹维劲　渠敬东　潘　涛

出版说明

自中西文明发生碰撞以来，百余年的中国现代文化建设即无可避免地担负起双重使命。梳理和探究西方文明的根源及脉络，已成为我们理解并提升自身要义的借镜，整理和传承中国文明的传统，更是我们实现并弘扬自身价值的根本。此二者的交汇，乃是塑造现代中国之精神品格的必由进路。世纪出版集团倾力编辑世纪人文系列丛书之宗旨亦在于此。

世纪人文系列丛书包涵“世纪文库”、“世纪前沿”、“袖珍经典”、“大学经典”及“开放人文”五个界面，各成系列，相得益彰。

“厘清西方思想脉络，更新中国学术传统”，为“世纪文库”之编辑指针。文库分为中西两大书系。中学书系由清末民初开始，全面整理中国近现代以来的学术著作，以期为今人反思现代中国的社会和精神处境铺建思考的进阶；西学书系旨在从西方文明的整体进程出发，系统译介自古希腊罗马以降的经典文献，借此展现西方思想传统的生发流变过程，从而为我们返回现代中国之核心问题奠定坚实的文本基础。与之呼应，“世纪前沿”着重关注二战以来全球范围内学术思想的重要论题与最新进展，展示各学科领域的新近成果和当代文化思潮演化的各种向度。“袖珍经典”则以相对简约的形式，收录名家大师们在体裁和风格上独具特色的经典作品，阐幽发微，意趣兼得。

遵循现代人文教育和公民教育的理念，秉承“通达民情，化育人心”的中国传统教育精神，“大学经典”依据中西文明传统的知识谱系及其价值内涵，将人类历史上具有人文内涵的经典作品编辑成为大学教育的基础读本，应时代所需，顺时势所趋，为塑造现代中国人的人文素养、公民意识和国家精神倾力尽心。“开放人文”旨在提供全景式的人文阅读平台，从文学、历史、艺术、科学等多个面向调动读者的阅读愉悦，寓学于乐，寓乐于心，为广大读者陶冶心性，培植情操。

“大学之道，在明明德，在新民，在止于至善”（《大学》）。温古知今，止于至善，是人类得以理解生命价值的人文情怀，亦是文明得以传承和发展的精神契机。欲实现中华民族的伟大复兴，必先培育中华民族的文化精神；由此，我们深知现代中国出版人的职责所在，以我之不懈努力，做一代又一代中国人的文化脊梁。

上海世纪出版集团
世纪人文系列丛书编辑委员会
2005年1月

不确定的科学与
不确定的世界

献给拉纳（Lana），约翰（John）和萨拉（Sara）……我生命中的至爱

对本书的评价

科学充满着不确定性。科学是如何发展，又是如何繁荣的？《不确定的科学与不确定的世界》在这方面给予外行人一个很好的了解机会。波拉克分析了社会不能或不愿处理全球环境问题的诡辩——常常借口没有足够的科学确定性对全球问题采取措施——然而同时保险业和其他行业只不过是按惯例预防未来的不确定性所造成的风险。此书写作清晰，没有辩论和公式，希望能够得到普通公众和政策制定者的广泛阅读。

——保罗·克鲁岑(Paul Crutzen)，

1995年诺贝尔化学奖获得者(因其在臭氧洞方面的工作)

人们对于科学有许多根本性和普遍性的误解。《不确定的科学与不确定的世界》一书能够澄清其中的一种误解——科学是精确和确定的。波拉克通过列举一些简短而清晰的事例，说明不确定性对于科学是常见的，瘫痪和无法行动的原因并不是不确定性，从而驳倒了“科学是确定的”神话。这本写作优秀的书有助

于减轻那些将科学不确定性滥用到制定政策并推进他们自己议程的特殊行业的误解。

——斯蒂芬·施奈德(Stephen Schneider),
斯坦福大学环境生物学教授,
《地球实验室:我们输不起的行星赌博》的作者

这本优秀的书将强有力地抨击两种危险的态度:一种是在用人们难以理解的概率进行描述的复杂世界中拼命寻找不可能的确定性;另一种是非常相信科学家是当今通过科学实验产生确定性的魔术师。波拉克生动而清晰地描述了诸如地球变暖等重大问题。阅读此书能够帮助我们在彼此冲突的“事实”面前作出很好的鉴别,这种能力对于21世纪是极为重要的。

——奥布里·曼宁(Aubrey Manning),
爱丁堡大学荣誉退休教授,
《动物行为研究》的作者

公共政策争辩经常陷入科学不确定性及其风险所造成的困境之中。其实,科学与人类经历的许多其他领域一样,对于不确定性和风险是无法避免的。波拉克通过列举日常生活的例子,以一种通俗而有趣的叙述,揭开了围绕着科学不确定性的神秘面纱。

——约翰·霍顿(John Houghton)爵士,
气候变化政府间专门委员会副主席,
《全球变暖》的作者

最后,我们对不确定性——这个令人困惑的主题有了可靠而科学的了解。当你读完这本书,也许会有一些主题,你对它们更加不确

定，但你理解了为什么。

——詹姆斯·特赖菲尔(James Trefil)，

乔治梅森大学物理学教授，

《一名城市科学家》的作者

科学家常常陷入技术术语的泥潭，而无法与普通公众成功地交流重要信息。在《不确定的科学与不确定的世界》中，波拉克运用平实的语言和极具吸引力的事例探索了科学和日常生活中的不确定性。

——尼尔·莱恩(Neal Lane)，

赖斯大学教授，

美国总统克林顿(Clinton)前科学顾问和美国国家科学基金会前会长

内容提要

为什么科学永远不能回答以下引人注目的重大问题:

地球正由于温室效应而变暖吗?

恐怖主义释放炭疽孢子意味着什么危险?

应该采取什么措施应对口蹄疫或疯牛病的爆发?

为什么我们不能预测地震的发生?

科学的不确定性令许多人困惑。当科学家之间具有不止一种答案而且这些答案彼此冲突时，困惑就发生了。《不确定的科学与不确定的世界》将帮助人们穿越矛盾和不确定性组成的迷宫。此书向人们介绍了不确定性在科学中发生的方式，科学家如何适应和利用不确定性以及科学家如何在不确定性面前得出结论，从而使读者根据自己的日常经验从自身角度自信地评价不确定性。

致　　谢

我非常感激我的妻子拉纳和儿子约翰，在我写作的几个阶段阅读了我的手稿，并提出了尖锐但极具洞察力的评论和建议。他们本身都是多产和出色的作者，家里有这样优秀的编辑使我受益匪浅。斯默登（Jason Smerdon）、基弗（Boris Kiefer）、查普曼（David Chapman）、艾萨克斯（Drew Isaacs）和凯恩（Gordon Kane）也阅读了全部或部分的手稿，在许多方面对我帮助很大。我也感谢剑桥大学编辑劳埃德（Matt Lloyd）早期的鼓励和后来重要的评论。当然，所有这些给予我帮助的读者对书中的错误以及尖锐的观点无需负责。

作者自我介绍

所有的书都免不了介绍作者的背景和经历，因此请允许我稍微介绍一下自己吧。

我出生在美国农业的中心地带——内布拉斯加州，并在那里接受教育。我母亲是一位传统的家庭主妇，我父亲在自家的农场里饲养家畜。在内布拉斯加州，儿时的我以为这个世界是由泥土组成的，肥沃富饶的土壤经过辛苦劳作之后可以长出能吃的好东西。1954年，18岁的我进入了离中部家乡很远的纽约北部的康奈尔大学。在康奈尔，有纽约州令人叹为观止的冰蚀指状湖群之一的卡尤加湖，一些小溪流蜿蜒流入卡尤加湖，基岩暴露于这些小溪流流过的大峡谷。在那里我知道了土壤只不过是层层岩石之上的薄薄一层，是真正的陆地；而在那些岩层之中是化石，记录着地球过去的生命。我对其揭示的时间的广袤性深感敬畏，实际上，很快我就被地质学深深吸引了，并将其作为自己的职业选择。整个地球都是我的研究领域，而

且在某种意义上，我从来没回过家："一个在农场长大的人被外面的精彩世界深深吸引以后，怎么还愿意回到原来的小农场呢？"

我确实为了一个硕士学位(受到地质学方面出色的教师和每学期90美元低廉学费的吸引)回到内布拉斯加大学，随后继续在密歇根大学攻读博士学位，在哈佛大学做博士后。在密歇根大学与哈佛大学学习之间，我与一位名叫拉纳·舍恩伯格(Lana Schoenberger)的密歇根女孩结了婚。几年后我在密歇根大学看到一个招聘教师的广告，于是就去面试，并且非常高兴地得到了这份工作。从此我和拉纳就在密歇根大学的本部安阿伯安了家。我们有两个孩子，萨拉和约翰，萨拉在14岁时由于意外事故去世了，约翰是一位作家，住在华盛顿特区。我们全家曾两次旅居海外，1970～1971年住在赞比亚，1977～1978年住在英国。

在密歇根大学，我讲授过各种层次的课程，从为那些不是以培养科学家为目标的本科生开设的入门性地球科学课程到为研究生开设的专门讨论班。实际上，我也讲过各种类型的课：讲座、实验课、研讨班和野外实践。我们系与其他系不同，入门课程的任务分给了更有经验的教师，结果我给那些不会从事科学职业的本科生讲的课多了起来。这种课程的挑战就是培养学生形成一种意识，即科学与他们生活的联系，并使之领会"科学方法"的优势和弱点。我曾教过地质学一般性的入门课程，"地质学101"，以及其他名为国家公园地质概况、气候与人类、行星的地质概况与气候、全球变暖的科学与策略等课程。

多年来我主要的研究方向是，研究地球内部的热能以及热能是

如何随着时间推移而消失的问题。地球的热能是推动板块构造和大陆漂移的巨大“发动机”的燃料，地震和火山是它们的副产品。塑造地球表面的那些大规模运动，实际上是使这颗行星慢慢冷却下来的地球内部运动的体现。多年来，我和我的学生在非洲、南美和美国对地球内部的热能损失进行了实地测量。当有人让我用简单术语描述那个过程时，我回答说我到全球许多非常遥远的地方去测量地球的温度。

在过去的10年里，我的地热研究有了一个新的方向。我和我的同事开始认识到，地壳外部几千英尺的温度变化线图是记录了这颗行星过去1000年气候变化的档案。基本的理论是，如果地球表面变暖（或就此而言变冷），地表下的岩石将会感觉到这一变暖并且将其记录下来。表面变化的持续时间越长，变暖渗透到地表以下就越深。在关于全球变暖和导致全球变暖的可能原因存在重要争论的背景下，地壳这些深度所包含的温度信息，能够让人们把工业时代和前工业时代的地球表面温度作一对比。上述对比为评估人类对气候变化施加影响的大小提供了一种依据。

在这期间的学术生涯中，我也有获得管理经验的机会。我曾做过一段时间的地质科学系主任，还以研究所副所长的身份参与过学院管理。在国家的科学团体内，我曾是美国国家科学基金会（US National Science Foundation）地球科学部的几个顾问小组的成员。这些小组负责评估来自全国各地为寻求资金支持的科学家们的研究项目。我还在美国关于全球和环境变化的地球物理学联合会委员会（American Geophysical Union's Committee on Global and Environmental

Change)工作过4年，我的任务之一就是负责准备这个专业机构在全球气候变化问题上的状况报表。

我所在的大学就像许多其他大学一样，教师除了主要从事教学和研究以外，还要参加大学、团体、州和国家的各种活动。在我所在的系，我负责校友之间相互联系的事宜。除了有校友向本系学生和教师捐款这样明显的善举以外，还有更多来自于工作日之外与毕业生定期交流的微妙之处，这些人对其职业生涯中所面临的环境和约束、问题和解决方案有一种意识，他们对诸如水的质量标准、环境清理以及全球气候变化这些以科学为基础的问题有自己的领悟。这些交流有助于学术界人士评估他们的课程和项目的相关性和效用。

在过去的10年中，我有很多机会与大学以外的许多团体一起讨论诸如全球变暖、放射性废料的处理和地震预测之类的科学话题。参加这些讨论的人一般是没有科学背景、有思想又受过良好教育的成熟人士。这些讨论发生的地点各不相同，但是包括与来自全国和世界各地的密歇根大学的校友们的会晤，对扶轮社(Rotary)* 和基瓦尼俱乐部(Kiwanis)** 等团体服务组织的谈话，与专业记者进行的研讨会，与生态旅行者一起进行的南极探险，在美国参议院所做的证词，在白宫做的简要汇报，为州或联邦政府立法成员举行的研讨班，在报

---

* 扶轮社是个国际性的企业家组织，最早建于20世纪初的芝加哥。那时美国社会崇尚金钱万能，“孔方兄”差不多成为所有人追求的目标，大家都不择手段，商业道德败坏。美国青年律师哈里斯有感世风日下，便与三个朋友策划成立一个组织。哈里斯的最初目的是为了促进会员间相互信任，公正地进行各种商业活动，让人们能从商场上淡漠的合作关系变成亲密的伙伴。四人约定组织的聚会在彼此的事务所轮流举行，于是就有了“Rotary”(扶轮、轮流)这个名字。——译者

** 美国工商业人士的一个俱乐部。——译者

纸、电台和电视台上进行的访谈，以及在公共电台进行的电话采访。通过这些丰富多彩的活动，我大体了解了来自社会各个阶层受到良好教育的人们对科学事业所持有的诸多误解。也许这些误解中排在最前头的就是不确定性(uncertainty)概念。

科学的不确定性(scientific uncertainty)之所以令许多人困惑，并不是因为“科学家们不能回答所有紧迫的问题”这一事实让他们很难接受，而是因为当科学家们给出不止一个答案而这些答案又彼此冲突时，困惑就发生了。我发现我作为教师所作的主要贡献之一，就是帮助人们穿越矛盾和不确定性组成的迷宫。通过让人们了解不确定性在科学中发生的方式、科学家是如何适应并利用不确定性，以及科学家们如何在不确定性面前得出结论，我能够让他们根据自己的经历，从自己的观点对不确定性做出非常自信的估计。我希望有更广泛的人能够分享这些观点，所以写就了本书。

亨利·N·波拉克

安阿伯，密歇根

目录

1 对本书的评价

4 内容提要

5 致谢

6 作者自我介绍

---

1 第一章 奠定基础

5 第二章 科学的不确定

23 第三章 媒体能帮忙吗?

42 第四章 不熟悉导致不确定

61 第五章 发烧还是着凉?

84 第六章 对半的机会

100 第七章 我不太确定这是如何工作的……

121 第八章 如果……让我们来看看会发生什么

141 第九章 重建过去

163 第十章 预测未来

182 第十一章 完全出于意外

205 第十二章 在一种不确定性的氛围里

228 译后记

第一章　奠定基础

这是一本关于不确定性，特别是与科学相联系的不确定性的书。多少年来，科学不确定性一直受到自然科学家、工程师、医学研究者、社会科学家和哲学家的关注。但是对于从小短文到学术专著所展示的所有观点来说，科学不确定性的丰富性，常常没有得到那些不是专门从事科学的普通民众的重视，而且(或者)经常被他们所误解。

当然，不确定性并不局限于科学世界。它也是日常生活常见的事实。我们经常会以多种方式面对不确定性。今天会下雨吗？多萝西(Dorothy)姑妈乘坐的飞机会按时到达吗？股市会下跌吗？上下班高峰时发生一个事故会让高速公路混乱吗？这些日常的不确定性来回穿梭，我们继续前行，有时为它们做好了准备，但更多的是在其中艰难行进。

但是，不确定性也影响长期关怀。我的退休金够不够我和妻子在今后的20年里过上我们希望的丰足而舒适的生活呢？我们的健康状况允许我们在未来的30年内以一种自由独立的生活方式生活吗？

这些长期性问题很难回答，都掩藏在更大的不确定性之中。由于我们仅能选择一种生活，因此不可能回到“过去”走另外一条路。无法避免的事情是我们必须制定计划，做出决定，尽自己最大努力，始终根据当时最好的判断对我们的行动作评估和中期方向修正。

不确定性绝不仅仅局限于将来；它也具有过去知识的特征。收养的孩子想知道他们的亲生父母，家人很难重建让他们曾祖父母迁移出去的环境。军事历史学家继续重建戈登(Gordon)* 将军在喀土穆**的最后岁月或者卡斯特(Custer)少校最后矗立于山头俯瞰小加拿大盘羊(Bighorn)***的各种场景。地质学家远不能确定冰河时代的原因，古生物学家仍然在为鸟的进化而争论不休。由于过去的记录是不完全的，某种程度上是不准确的，因此我们对过去的理解是不确定的。我们持有的证据经常是互相矛盾的。

纵观生活，人们陷入不确定性之中。他们会以合理的、可接受的和非敌意的反应例行公事地适应这些不确定性。在简单层面上，城市居民会有带把伞来应对下雨的可能；较复杂层面上，农民可能会参加商品期货市场防止受到可能的干旱的影响。退休金管理者在面对长期相当大的经济和政治不确定性时，会例行公事地做投资决定。房主和车主会买保险来预防在不可预料的将来发生大灾难。这些都是面对不确定性采取的理性措施。然而，就做决定者而言，在面对类似层面的不确定性时，他们有时不情愿采取措施处理以科学为基础的复杂问题，部分原因在于他们可能感到对随之而来的科学不确定性

---

* 戈登，全名为查理士·詹姆斯·戈登(Charles James Gordon)，苏丹总督。他曾被围困在苏丹城内，英勇无畏地与死亡和命运搏斗，坚守了一年之久，最后，终被砍死。在威斯敏斯特教堂内的纪念碑上，刻有这些发人深省的文字，“他把他的钱财捐赠给穷人，他的怜悯转向悲痛之人，他的生命献给国家，他的灵魂交给了神。”——译者

** 苏丹首都。——译者

*** 北美洲西部山区产的一种野羊。——译者

进行研究和评价还没有作好足够的准备。 全球气候变化的主题既说明了科学的复杂性(scientific complexities)和不确定性，也说明了在制定合理政策来处理地球气候变化的许多方面遇到了困难。

一些主题，无论是日常的还是科学的，会贯穿本书的各个章节，它们或多或少解释了我在适应不确定性方面的观点：

- 不确定性总是伴随我们，它决不可能从我们的生活(无论是个人还是作为社会整体)中完全消除。 由于不确定性的存在，我们对过去的理解和对未来的预测总是模模糊糊。
- 因为不确定性永远不会消失，对于未来的决定，无论其大小，总是在缺乏确定性的情况下做出的。 在做决定之前一直等到不确定性完全消除是对现状的含蓄支持，常常是维持现状的一个借口。
- 预测长期的未来是一件危险的事情，很少能做出与现实非常接近的预言。 随着未来的逐渐展开，需要做“中期方向修正”以便考虑到新信息和新进展。
- 不确定性，远非前进的障碍，它实际上是创造性的强烈刺激因素和重要组成部分。

不确定性花园

整本书将带着你进行一些科学远足，说明不确定性是如何被编进科学事业的结构之中。 这些艰难旅行，许多在我进行科学研究的领域——地球和环境科学之中。 特别是，对当代全球关注的话题——全球性气候变化在许多地方都要涉及。 在 20 世纪 90 年代也许没有哪一个话题能像全球气候变化一样一直吸引着人们的注意力，并且围绕

它产生如此激烈的争论。不同时期成为焦点的问题是，气候变化的现实性、原因、结果，以及因气候变化导致的政治反应、经济反应和社会反应。作为缓慢发展的复杂的全球性现象，总体上它展示了科学不确定性的许多迷人方面，说明了科学家是如何在不确定性环境下工作和发展的。

本书所列出的科学远足可能被认为是在“不确定性花园”中进行的游览，是对广袤而不规则的土地进行的探索。这片土地由生长着一年生植物和多年生植物的小块土地、一些新开垦的土地、稀有品种、杂草、灌木丛以及迷宫构成。这个花园中的每一个领域都揭示了不确定性的不同方面。对来自科学不确定性的每一次洞察来说，通常都会在科学王国之外发现具有启迪作用的经历，这种经历让读者意识到科学世界与他们自己的世界并非如此不同。的确，科学是我们生活的世界中一个重要的、可理解的又极具影响力的组成部分。

通过对科学和日常生活中存在的不确定性进行比较和类比，我的目标是让读者用与处理生活中的不确定性相类似的那种方式去理解和适应科学的不确定性。我希望读者在读完这本书之后有着这样的感觉：科学不确定性与人们在日常生活中遇到的并例行公事般适应的许多其他不确定性相比，并不会让人更加困惑或产生怀疑。随着对科学不确定性的更好理解，读者能够识破偶尔让科学对于社会问题的价值及其相关性模糊不清的疑云。在理解不确定性的进程中，对于回答“科学能够提供什么和不能提供什么”的问题，他们会更加充满自信。

第二章 科学的不确定

"科学"是日常生活之外的独立组成部分，我坚决向这种观念挑战。我们生活在一个科学的时代；然而我们却认为知识只是一小部分人的特权……这种观念是不正确的。科学的素材也是生活的素材。科学是活生生的现实，是我们经历的每一件事情是什么、怎么样和为什么的问题。

——雷切尔·卡森(Rachel Carson)*，在接受1952年国家图书奖(因《围绕我们的海洋》一书而获奖)时的演讲

科学，正如卡森所观察到的那样，是生活的一个组成部分。它具有自身的优势和弱点、成功和失败、疑问和不确定。当科学家试图理解一个细胞异常是如何引起癌症，一个基因是如何传递信息促进

* 卡森(1907～1964)，美国生物学家。主要著作有《在海风的吹拂下》、《围绕我们的海洋》、《寂静的春天》。真正让她获得广泛关注的是《寂静的春天》，它的诞生标志着现代环境运动的开始和环境新闻走向成熟。——译者

机体发育，一个生态系统是如何应对城市的拓展或者整个地球是如何应对大气中化学成分的长期改变时，这些研究的每一个阶段都被包围在不确定性之中。不确定性的产生有许多方式，它的本质也许会随时间而改变，但是搞科学的人决不可能游离于不确定性之外。

科学会因为不确定性而衰弱吗？恰恰相反，许多科学的成功正是由于科学家在追求知识的过程中学会了利用不确定性。不确定性非但不是阻碍科学前行的障碍，而且是推进科学进步的动力。科学是靠不确定性而繁荣的。遗传性状如何被复制的不确定性，最终导致了双螺旋分子结构的发现。的确，也许有人会争辩，是确定性而不是不确定性阻碍着科学。17世纪哥白尼(Copernicus)、开普勒(Kepler)、伽利略(Galileo)为了推翻地球是太阳系中心的观念所进行的长期斗争，[1]在当时盛行的“地球在宇宙结构中占据一个非常特殊的位置”的神学确定性面前，继续进行着。

实际上，科学家面对的不确定性同我们日常生活中遇到的不确定性没有如此大的不同。在许多文化中，冒险作为成功人士应该具有的品质受到人们的赞美。但是风险恰恰是因为不确定性而产生。在不确定性面前阐明风险、采取措施和接受风险的意愿和能力，被认为是一种特别的优势。虽然有一些冒险后来被证明是不明智的，但是没有冒险，就是对现状的含蓄接受。不愿意受不确定性的激励，才是前进的真正障碍。

具有讽刺意味的是，不是科学家的那些人经常将科学等同于确定性，而不是等同于不确定性。他们要求对以下事情进行高度精确和准确的预测：日食，海洋潮汐的日常进展，当地日出日落的准确时间，在一个遥远的行星上着陆的航天器内时钟的精确度。确定性的另一个方面是和技术可靠性相关的，当人们拿起电话，打开电视，或

者转动轿车的点火钥匙时，都期待着设备正常运转。确实，当事物不是按期望或预料的发生时，通常会出现某种程度的惊奇和不满。绝大多数人都不喜欢意外，而且对不可预测性（unpredictability）和不确定性感到某种程度的不适。

其他背景下的确定性是满意的一个来源。让信徒对死后生活放心的宗教信念可减轻对死后进入地狱的担忧。诸如“较小的政府是更好的政府”或者“没有比税收更好的事情”这样的政治颂歌，可将这些人从评估广泛的公共政策问题的负担中解脱出来。如果将一个充满灰色阴影的世界重塑成一个只是由黑色和白色组成的更简单、更坚实的实体，就会消除那种掂量细微差别的艰巨任务，取而代之的是确定性提供的舒适性。

在对复杂自然系统的理解中，当科学家不能说明高水平的确定性时，有时在普通公众中会有一种不耐烦和不满的潜在倾向。2001年末，在美国的政府大楼和邮政设施中出现了以炭疽孢子（anthrax spores）形式出现的生物恐怖主义（bioterrorism）。然而，一段时期以来，在公共健康团体和国家疾病控制中心内，存在着对以下问题的不确定和困惑：炭疽病是如何被传播的？多高的孢子浓度会有危险？炭疽孢子如何引起身体虚弱？公众想要知道答案，公共健康从业者却不能立即提供。与此类似，在英国，2001年口蹄疫的爆发遇到一系列关于如何遏制这种疾病的科学观点。采取的遏制策略是大量宰杀周围的畜群，但是科学观点远未达成一致。疾病得到控制以后的很长时间，关于宰杀策略是否必要或有效的争论还在继续着。

当科学家承认，他们不知道诸如“疾病在一个生态系统传播”这样一种复杂自然现象的所有方面时，公众有时会将此理解为科学家对

这一事件一无所知。这样就会导致公众对科学共同体能力的不信任，而不信任的副产品是普通公众更加频繁地对一些狂人、吹牛者和十足怀疑论者所发布的毫无价值的谬论怀有兴趣。在科学权威和确定性的氛围下，这些伪科学家(pseudo-scientists)做出了一些从未经过严格检验的断言，而严格检验正是真科学的基础。

算命者(fortune-tellers)、看手相者(palm readers)、透视者(clairvoyants)、占星学家(astrologers)(名单可能会继续延长)，所有这些人都是靠不让或不愿意让他们的顾客辨别出这些实践完全缺乏逻辑基础和科学观察而发家的。无论如何，绝对没有任何东西会提供给这些骗子以使其获得信任。但是他们的宣言总是经过精心构思以便给顾客留下这样的印象——有一种非凡的力量在客观地运行着。在下一章，我将描述有关这方面一个特别让人吃惊的例子，太多的对地震理应更加了解的人对大地震所做的预测是过于严重了。

当然，有一些严肃的学者对以下观念提出挑战：科学是通向宇宙真理的唯一途径。有一个哲学流派，笼统称呼的话可称为后现代主义(postmodernism)，对“科学家是中立的和客观的”和“科学知识是毫无偏见的理性思考的结果”这些观点表示怀疑。极端的情况是，它会质问在人类建立的智力结构之外是否存在着一个确定的自然世界。这种边缘性观点把科学视为由科学家创建的拥有一套规则的游戏，并声称如果我们不接受这些科学游戏的规则，那么在对自然界的理解中，科学显而易见的成功将无法维护。此立场的一个子题目就是科学是一个自我服务的概念和实体。

1996年，当纽约大学的物理学教授索科尔(Alan Sokal)博士，向一家被认为是支持这一特别哲学的杂志提交一篇文章[2]以供发表时，后现代主义的观点成为引人注目的焦点并被大加嘲讽。这份稿件的

标题为《超越边界：针对量子引力的变形解释学》，这篇文章似乎是在传递一种后现代的味道。由于一名物理学家已经提交了手稿，杂志社的编辑们欢迎给似乎能够从内部腐蚀科学基础的科学家发表文章的机会。但是索科尔的文章是一匹特洛伊木马，是经聪明构思的骗局，本文说明的并不是科学方法在哲学上的脆弱性，而是表明编辑们很容易上当受骗。索科尔运用绕来弯去的语言和结构，写了一篇看似博学的论文，其实只是掩藏在伪科学行话(pseudo-scientific jargon)之下的胡言乱语。过分渴望的编辑们上了圈套，发表了索科尔的文章。一旦文章印了出来，索科尔揭露了这个骗局。“索科尔事件”(l'affaire Sokal)就像它的称号一样，含义有许多，但就我的目的而言，主要的观点是这样：有些人，无论其受过教育与否，他们只是相信科学不提供任何特殊的东西。他们怀疑或者完全忽略科学结果。

还有另一种类型的人，除非科学同他们挚爱的其他信仰发生了冲突，一般情况下他们会接受科学结果。就在写作本书时，我读了国际平坦地球研究协会(International Flat Earth Research Society)主席约翰逊(Charles K. Johnson)的讣告。[3]除了行星形状这一特殊的困扰之外，约翰逊先生看起来就像飞机机械师一样过着相当正常的生活。除了有关地球的形状之外，他很少有与科学共同体不一致的意见。他这样解释“阿波罗号”宇宙飞船的宇航员从月球上拍摄的地球照片：月球登陆不过是精心策划的骗局，照片只是这一骗局中的一个道具。我们也许会对这一怪论一笑置之，但是科学不确定性的水池，却由每一位成功关注从另类角度观察自然界的约翰逊一步步加深了。1994 年，一份民意测验[4]显示差不多有 1/10 的美国人认为月球登陆是伪造的。好莱坞在诸如 1998 年拍

摄的电影《做大英雄》*等影片中也添油加醋。在这部影片中，美国总统通过编造对阿尔巴尼亚的虚假战争，包括用描绘破坏和屠杀的虚假电影脚本在内的虚假入侵的方式，试图将注意力从个人的困境中转移出去。

在科学和个人信仰之间一个更广为人知的冲突，集中在《创世记》中对上帝创造万物的《圣经》的描述上。这一争端是关于《圣经》上的每一个字是否都是正确的。上帝是否在仅仅6天之内创造了整个宇宙以及其中的每一个生命？地质学家和进化生物学家提出了很有说服力的理由：并不是所有的现代生命形式在地球诞生之日就出现了，今天绝大多数的生命都是经漫长的地质年代由其他生命形式进化而来的。但是拘泥于对《圣经》字面解释的人并不接受对《创世记》哪怕是极小的偏离。如果《创世记》的每一个字都是正确的，那么现代地质学和生物学必定是错误的。

神创论者(creationists)目前正从事着以下事情：通过证明一种"创世科学"(creation science)的努力，证明进化生物学的原则是不正确的。这种所谓的"创世科学家"试图找出进化生物学的逻辑缺陷或观察错误来"否证"它。然而，他们还没有运用同样的精力来检验《创世记》中建立的假说。他们甚至不愿意承认《创世记》中的叙述是一个假说，更不要说检验了。他们无法想象运用实验和观察否证《创世记》，因为"创世科学"的从业者不是真正的科学家。神创论者在他们基本的立场面前决不会退让，即所有活的生物都是上帝直接同步创造的。他们不允许出现这样的可能性：《圣经》中关

* 此电影的剧情如下：在新一轮大选开始前两周，现任美国总统遇到了麻烦。新闻媒体大肆攻击他在白宫召见一位专为名人物色美女的皮条客。总统的幕后人马没一个关心这事是真是假，反正支持率肯定会下降，现在他们关心怎么把支持率再搞上去。——译者

于创世的描述是不正确的或有朝一日被证明是站不住脚的。相比较而言，真正的科学从业者，很容易承认不确定性，而且在不确定的环境中工作感觉如鱼得水。在真实科学（real science）中，很少有概念被认为是绝对正确或被绝对确定地接受。

的确，真科学（genuine science）是建立在假设的基础之上，这个假设中的概念有可能是错误的。当一个概念被证明在逻辑上不一致或与直接的观察相反时，证伪（falsification）就发生了。考德威尔（Lynton Caldwell）在对齐默尔曼（Michael Zimmerman）的《科学、非科学和谬论》[5]一书的评论中，将科学描述为“从可能的正确中分离出可论证的错误”[6]的过程。仅仅是谬误能被证明，而真理是不能被证明的，这是科学的根本基础。真理仅仅是多次努力后未被切割的幸存者。事实上，科学部分是通过“不断探索概念的最薄弱部位”这样的方式获得进展的，而这些概念在解释某些自然现象时取得了部分成功。这种对弱点的无止境的研究可能会揭示理论的微妙矛盾，最终导致对最初概念进行修改或抛弃。

同行评议

这些科学对抗发生的环境是什么？核心的过程叫做同行评议（peer review）。当科学家希望告诉世界他们从事的某些研究时，应当遵循已制定的路线。通常情况下，第一步是要在专业学会会议上做有关研究的口头陈述。这需要他们将自己的贡献做成一简短的书面摘要，并提前交给为会议制定议程的委员会。接着，将这一摘要在议程中公布出来，以便其他人决定自己是否要出席。在会议上，比较典型的做法是，作者要对其工作做 10 到 20 分钟的口头陈述，在此之后，会有听众提问或讨论。在讨论中表达的观点范围很广：同

意、不同意、怀疑、赞美、嘲笑等等。

如果科学家的口头陈述受到充分鼓励，那么他或她接着要为其工作准备一份比较长的书面报告，交给学术期刊以待发表。接下来编辑会将手稿送交在所交论文的大致领域工作的其他科学家，征求他们对此项研究是否适合发表的意见。要求同行评议者从几个方面评估这项研究。该项研究是否具有新颖性和独创性？所运用的方法是否与研究目的相匹配？实验设计或理论推导中有错误吗？结论是从观察和数据中直接得出来的吗？结果具有的不确定性水平是什么？让有经验的从业者对研究进行检验是杜绝有错误研究的过滤器，但是允许满足一定标准的研究成果发表以供其他读者阅读、评估、竞争或复制。事实上学术期刊发表的所有论文都要经过同行评议的检验。

同行评议过程并非没有错误，但是经同行评议过滤掉的有缺陷科学远比其偶尔的失败要多。有时同行评议对后来被证明是不正确的特别重要的主张虽然持有怀疑，但考虑到自我修正的过程，仍然做出有利或较宽大的决定。1999 年，一组物理学家为一种新的超重元素的存在提供了实验证据，该元素在元素周期表中编号为 118 号。然而其他实验室的实验人员，包括最初那组实验人员，尝试通过重复实验重复实验结果，但都没有成功。经过两年的失败之后，最初的研究小组宣布撤销他们的声称，承认他们可能是曲解了第一次实验的数据。2002 年初又发生了一次类似事件，有一篇发表在一家颇有声望的杂志上的论文宣称观察到了核聚变的证据，即先是有小泡形成，接着受声波激发，在有机溶剂中发生内爆。[7]在同行评议过程中，这篇论文非常具有争议性，但是由于实验的结果一旦是真的，就会有着非同寻常的意义，因此编辑决定发表它。毫无疑问，许多其他实验室的科学家为了要证实所报告的结果或想证明此结果无效，将会对此实

验进行重复。

因特网作为一种新的媒体，对同行评议提出了重大的挑战。任何拥有计算机的人，都能将他或她的研究，无论是合理的还是有缺陷的，相关的或不相关的，重要的或微不足道的，放入公众领域让任何人阅读。这种现象大大增加了此项研究的用户进行检验和评估的负担。同行评议的角色在于过滤掉有缺陷的研究并防止其发表在科学杂志上，现在这种守门员的角色(gate-keeping role)变成了让所有的单个读者在因特网上冲浪。因特网是很卓越的平均主义者，因为任何人几乎能邮寄任何东西，但是确定邮寄的东西是否真实或有价值则是个人用户的事情。在因特网发展之前的早些时期，不经过编辑审查就让自己的思想或观点公诸于众仅仅是某些有钱人的特权，他们能够通过籍籍无名的出版社自掏腰包出版。将评估的责任转交给每一个浏览万维网的个体，使得公众理解科学和不确定性更加必不可少。

不确定性的传播者

不喜欢科学所述东西的那些人，常常准备对科学发动微妙或并不微妙的攻击。当他们发现某种特定研究结果具有威胁性时就实行这种形式的攻击。他们常常争辩道，如果科学是经“合适地”构想或执行的话，就会出现一个不同的结果(暗示着他们喜欢不止一个的含义)。对科学可信度进行这种特别攻击的代码，常常出现的有“不健全科学”(unsound science)、“未决定科学”(unsettled science)、“不确定科学”(uncertain science)、“拙劣的科学”(poor science)、“垃圾科学”(junk science)等类似话语。与同行评议层次的批评不同，这些批评总是发生在平常的科学渠道或科学标准之外。它们通过付广告费出现在报纸上，或出现在致编辑的信件中，或通过参加电

台或电视台的访谈节目来表现。

这些代码描述被石油和煤炭工业定期用于有关全球气候变化的原因和结果的评论。在刊登广告的醒目位置，埃克森美孚公司(ExxonMobil Corporation)[8]频繁贬低证明气候变化或提供“使用化石燃料[9]会引起气候变化”这样证据的科学研究。作为最大的国际石油公司之一，埃克森美孚公司在使用化石燃料方面有很大的利益，由此它试图推迟立法或让限制二氧化碳或其他“温室气体”(greenhouse gases)排放到大气中的国际条约出轨。人们可以想象，如果化石燃料工业有检验温室气体在气候变化中作用的重要观察资料，那么他们将会战斗在科学的竞技场——常常发生科学争论的受到同行评议的科学杂志，而不是战斗在媒体里或街道上。但是化石燃料工业对赢得伦敦、柏林、华盛顿等国家首都的政治斗争更感兴趣；他们将时间浪费在公众竞技场，迷惑并逐渐破坏公众对科学结果的信任。

通过散布对科学的怀疑和不确定性以对决策人产生重要影响的策略并非没有成效。2001年3月，惠特曼(Christine Todd Whitman)，布什(George W. Bush)政府新任命的环境保护局局长，取消了前政府公布的对饮用水里砷含量的更严格限制，并开始重新评估建立那些规则的科学基础。“我们将会运用强大的科学……确定新的限制应该是什么。”[10]这种陈述的目的只有一个：逐渐破坏公众对以前科学研究的信任。而刚刚废止的规则就建立在这些科学研究的基础之上。在制定那些规则之前，评估环境中砷的科学、砷对公众健康的影响，以及美国国家科学院关于砷的检测报告已有超过10年的历史。这显然不足以战胜以下两方面的反对意见：一是采矿工业，在矿石加工中必然排放砷这种副产品；二是社区，人们若想继续饮用更好的水，就

必须改进净化装置。当然环境保护局仅对白宫的立场做出回应。2001年4月24日，布什总统在白宫国宴厅对一群环境青年奖获得者说："我们将要在健全科学的基础上做决定，而不是根据周围的时尚或听起来可能很好的东西做决定。"[11]6个月之后，国家科学院在政府的要求下对相关证据作了评估之后，确定对砷实行新的较低含量限制是完全恰当的，实际上也许还不够严格。当然，没有新的科学数据，也没有另外的"健全科学"出现支持这种含义，即先前宣布的修订案是建立在不健全的科学基础之上。

对科学的这些攻击几乎不是什么新现象。1952年，当卡森宣布杀虫剂DDT的广泛使用对鸟类繁殖具有毁灭性的影响时，[12]杀虫剂工业将其见解当作建立在脆弱科学基础上加以嘲笑。数十年来，烟草行业一直否认存在任何证明吸烟有害健康的科学证据。20世纪70年代，人们开始认识到环境中的铅会对健康产生危害，产品中含铅的汽油生产者却嘲笑这一科学。在位于美国东北部的那些州，当人们发现酸雨是由中西部的发电厂燃烧高硫含量煤引起的时候，电力工业对这项研究不屑一顾。20世纪80年代，当遇到氯氟碳化合物(CFCs)正在破坏平流层*臭氧这样的论断时，化学工业争辩道论断后面的科学是脆弱的和不确定的：

> 国际化学工业坚决否认臭氧层状况和日益增长的CFCs的销量之间存在任何关联。为了对此理论进行质疑，工业界迅速开

---

* 平流层，英文为stratosphere，也译作同温层。包围着地球的那层空气叫做"大气层"或简称为"大气"。大气的密度随着高度而减小，最后和星际气体连接，所以，它们之间不存在一个截然的界限。以气温变化为基准可将大气分为对流层(变温层)、平流层(同温层)、中间层、暖层(电离层)和散逸层等五层。平流层位于对流层之上，顶端离地面大约30公里，其特点是：温度大体不变，平均在-56.5℃左右，几乎不存在水蒸气，所以没有云、雨、雾、雪等天气现象，只有水平方向的风，没有空气的上下对流。——译者

展他们自己的研究和公共关系。[13]

当公共关系运动迷惑民众的时候，科学仍然坚定着自己的立场。1995年，诺贝尔化学奖被授予罗兰(Sherwood Rowland)、莫利纳(Mario Molina)和克鲁岑，表彰他们阐明了CFCs引起臭氧损耗的机制。这是诺贝尔奖第一次也是唯一一次对环境化学研究予以承认。

所有经过同行评议的科学研究都会被认可为伟大科学吗？当然不是。我定期阅读科学杂志，提交供发表的研究报告，也为他们做同行评议。绝大多数科学家会承认，杂志除了有大量重要的研究结果外，还包含一些正确但价值不大的稿件和一些后来被证明在方法上有缺陷的其他稿件。偶尔(但是会很少)有一次欺骗性的投稿，描述从来没有完成的工作或子虚乌有的结果，当时会成功地溜掉，直至后来有人足够仔细开始怀疑并且检查它时，才揭开它的假面具。当然，科学家不想让不可靠、不细致、没有价值的科学搅乱杂志和扰乱知识体系。但是*不确定的*科学，*未决定的*科学，几乎完全不同于不可靠的科学。科学中事情的正常状态是未决定的和不确定的，没有哪项新研究会完全消除不确定性。当较早的问题被回答时，新的问题又出现了。为了避免这听起来像无用踏车，我向你保证它肯定不是。不确定性根本不会让你失望或无力，相反它是科学家力量和创造性的源泉。不确定性是一种挑战，是科学进步的催化剂。

“人们热爱科学，他们只是不理解它”

为什么许多人在接受和适应科学不确定性方面都有一个如此艰苦

的时期呢？有比宗教信仰给予安慰的确定性、骗子提供的显著确定性或一些工业为了维护它们的经济利益而散播的让人迷惑的烟幕这些原因更深层次的理由吗？ 我认为，发生此类问题在更大程度上取决于以下事实：通常绝大多数人缺乏对科学的初步理解。 这种科学文盲(scientific illiteracy)为确定性的呼吁和不确定性的混乱提供了生根发芽的肥沃土壤。

C·P·斯诺(C. P. Snow)* 在其著名的《两种文化》**[14] 一书中，概略说明了在现代大学和社会中普遍存在的科学学科和人文学科相互分离的理解性鸿沟。 这一观点隐含在本章开头卡森的引文中。 1996 年，有一篇讲述美国科学教育的文章[15]是这样讲述的："美国人热爱科学，他们只是不理解它。" 的确，时常会听到这样悲观的论调：普通公众从来就不理解科学，更不要说领会不确定性的奥妙了。 如果我们不是在科学理解的如此之薄的冰层上行走，我们还会那么容易遭受狂人宣言或者特殊行业施放的迷惑烟幕的攻击吗？如果我们对科学不是如此陌生，也许就不会如此轻易受到这种迷惑的捕获。

理解科学的这些问题很早就始于教育系统的不完备。 在一种非常重要的意义上，孩子们天生就是自然科学家。 他们沉浸在奇特的世界里，对他们周围的一切事物充满好奇。 他们看，他们摸、听、闻和尝。 他们观察这个新的世界，并且处理和评估从每个方向到达他们的信息流。 他们探索、实验，从自身的错误中学习。 接下来他们就去上学了。

---

* C·P·斯诺(1905～1980)，英国作家、物理学家。 他的著作包括长达 11 卷的《陌生人和兄弟》、《两种文化及科学革命》和《现实主义者》。 ——译者

** 参阅 C·P·斯诺：《两种文化》，纪树立译，北京：生活·读书·新知三联书店，1994 年 3 月第一版。 ——译者

美国的学校教育，至少就科学调查来说，向孩子们介绍了一种新的方法论。这种新的方法论强调科学并不是像当初学步的孩子经历的好奇和探险的继续。相反，学校的科学不是作为探究的过程，而是更频繁地作为成就的复述出现。事实是最重要的。学生被告知：世界是圆的；地球绕太阳旋转；一年有365天；昆虫有6条腿；亚马孙河是世界上最大的河；珠穆朗玛峰是最高的山峰；岩石可以被分成火成岩、变质岩和沉积岩等种类；原子包含质子、中子和电子。新的强调是用人们确认的每一个受到良好教育的人必须知道的“事实”填满小小的脑袋。科学展现的是答案而不是问题。降级到这种遥远的背景，是对“事实”如何被确定、“事实”可能是多么的持久或短暂和我们相信这些“事实”是多么的确定或不确定等进行探究的过程。

答案，就像我早先提到的一样，对一些人来说比问题更让人受到鼓舞。在简单背景下，不确定性可能被解释为“它可能是这样或可能是那样”，但是这样的看法常常被认为是非常危险地接近文化和道德相对主义的深渊。在这种相对主义里，在正确与错误之间有一道灰色的阴影会把年轻人引入迷途。“不”，文化绝对主义者说，“有一些事情我们是确定的，我们不能试图将这个问题和模糊真理的不确定性相混淆。”

考试，考试……

学校教育的成功，常常是通过“对学生实行地方性和全国性的标准化考试来评价他们在阅读、数学和科学方面的成就”这样的方式来衡量的。一些学校系统和它们的老师被按照学生在标准化考试中的成败划分等级。把政府对学校的拨款与标准化考试的成绩联系起来

的提议，被执行得有条不紊。一些学校认识到他们的政治支持，甚至财政支持可能取决于考试成绩，因此他们“为了考试而教书”就不足为奇了。然而，从国际性眼光来看，即使是为了考试而教书并没有让美国的中学产生引人注目的成果。1999年，第三届国际数学和科学竞赛，对23个国家的学生进行了测试，结果显示美国8年级学生(13岁学生)的科学和数学成绩在国际平均线以下。[16]就美国国内而言，国家教育进步评估机构每隔几年要对4年级、8年级和12年级学生(分别是9岁、13岁和17岁)进行一次数学测试，结果划分为“初级以下”、“初级”、“熟练”和“高级”几个等级。2000年举行的测试中，9岁和13岁的孩子中仅有三分之一、17岁的孩子中不足五分之一达到熟练水平。[17]

当然，我承认科学进步，实际上任何生活努力中的进步，一定有着包括基本的文字能力和数字能力的教育基础。阅读、书写和运算能力，确实成为生活要素的一部分。但是，就必要性来说，如果仅仅将它们作为教育的目标，我们将会欺骗学生和社会。其他重要的技能，例如如何仔细观察、如何批判性地思考、如何处理冲突、如何发展团队精神，都是难以被测试的，但是对于学生的成功和学生整体来说，可证明是同等重要的或者是更加重要的。

这种对获得“知识”的强调，在整个初等和中等教育体系中持续存在，在许多高等教育课程中也未被削弱。用于对科学做引导性纵览的课程的教科书，太多的是对我们确实知道的知识的一种无趣的概述，对于我们不知道的知识却没有做一个激励性的总结。在这些教科书中对科学前沿的描述在哪里呢？什么是没有回答的问题？这些问题可能会让富有想象力的学生兴奋起来，重新点燃他们在孩提时代曾经拥有的对自然的好奇。为什么他们没有了解到那个领域具有的

不确定性？

科学随时间如何进化的历史揭示了科学家在过去的时间里有一个不成功的开始，是沿着死胡同前进的。马查多（Antonio Machado）*，20 世纪早期西班牙的一名诗人，表达了这种思想："对于行人来说，是没有路的。当你行走时你就创造出了路。"一个学科的历史，虽然通常没有在包含此领域的不确定性方面进行表达，但是它能够突出某一特定时期传统的知识，显示在冲突性的观察和竞争性的思想面前，传统知识如何被拆散并被较新的概念替代。对时间和历史的透视为了解以下问题提供了机会：科学作为探究的领域是如何发展的？探究性的问题和批判性的思考是如何促进人们更好地理解的？没有任何历史的背景，学生们必定只是满足于快速得到今天的答案，而不是对昨天或明天的问题进行探究。

在研究生学位计划中，原则上主要是训练职业科学家和未来的教授，然而大学却在尽力取消这一切。学生进入研究生院以后，长时间集中在给出答案而不是提出问题。他们对阐述一个研究项目感到困难，而此研究项目能够提出一个有趣而重要的问题，并且能够设计一条解决问题的大致通道。在我所在的大学里，许多科学界的同事并不乐于帮助学生摆脱束缚学生好奇心的教育限制。虽然许多系的教师自身对研究作过想象，但更多的是把研究生视作他们个人研究机器的嵌齿。他们没有让学生构想一个研究问题并提出解决这一问题的方法。相反，学生经常是被分配一个范围缩小的项目，此项目很少放在更大的研究视野下进行解释。他们被指导如何使用最复杂的

---

* 马查多（1875～1939），西班牙诗人。出版的诗集有《孤独、长廊和其他诗》、《卡斯蒂利亚的田野》等。——译者

研究设备测这测那，但就测量相关性来讲，经常被认为是不重要的而不加考虑。他们学到了知其然，却很少学到知其所以然。当然，当他们获得的技能随着下一代仪器的使用而变得过时的时候，许多人醒悟到他们已背离了科学。

因此为什么我们有这么多的科学文盲？为什么人们容易受过分简单的思想和虚假宣言的影响？为什么他们对科学不确定性感到迷惑？我认为，部分原因在于大多数学生接受的科学教育抑制了他们的自然科学直觉。许多学生在小学和初中就对科学丧失了兴趣，原因在于没有利用他们对自然的好奇。接着高等教育系统让这一问题延续下来，主修科学专业的毕业生是吸收了所有的“事实”而不是准备向这些事实挑战。同样愚蠢的体制用于培训新的小学和初中教师，这些受培训的教师再重复这些模式。研究生阶段，培训出的新科学家更多的是有能力的执行者而不是富有想象力的设计师。简而言之，许多国家共有的教育实践已经导致了成年人普遍对科学满怀兴趣同时又对科学感到迷惑，主要原因是他们不理解科学家是如何着手做提出问题和评估答案这样的事情。

继续不确定性花园的比喻，我们已经开始了在东方帐篷的旅行，那里的展览已经说明了科学和不确定性的一些社会学的、政治的和教育的方面。下一章仍旧固定在东方帐篷中，在那里我们将要集中到一个位于科学和大众之间的特殊机构——大众传媒。媒体有助于向大众传播科学吗？不仅仅是根据科学的成功、成就和确定性，而是将其作为一种探究过程或探究方法来传播吗？这种探究过程或探究方法是受失败刺激，而且是在不确定性投射的阴影的微弱暗淡的光亮中繁荣发展的。

注 释:

1. 这段历史将在第六章详细描述。

2. Alan D. Sokal, Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity, *Social Text*, 1996.

3. *New York Times*, 25 March 2001.

4. Marc Fisher, *Washington Post*, 20 July 1994.

5. Johns Hopkins University Press, Baltimore, MD, 1995.

6. *The Environment*, vol. 38, n. 6, p. 25, 1996.

7. Taleyarkhan, R. P. et al., Evidence for nuclear emissions during acoustic cavitation. *Science* vol. 295, pp. 1868～1873, 2002.

8. 范例参见 the *New York Times* for 23 March, 10 August, and 21 September 2000。

9. 煤、石油、天然气等燃料被称为化石燃料，它们是在很久前通过地质过程形成的。这些燃料存在于组成地球的地壳的岩石中，通过在地表开矿或抽吸获取。

10. *New York Times*, 21 March 2001.

11. http://www.whitehouse.gov/news/releases/2001/04/20010424-1.html.

12. *Silent Spring*, Houghton Mifflin, New York, 1952.

13. 注释参考 Dotto and Schiff, *The Ozone War*, pp. 149～165, by Richard Elliot Benedick in *Ozone Diplomacy*, p. 12, Harvard University Press, Cambridge, 1991。

14. *Two Cultures and the Scientific Revolution*, Cambridge University Press, 58 pp., 1963.

15. Michael Carlowicz, *EOS Transactions of the American Geophysical Union*, 27 August 1996.

16. http://ustimss.msu.edu/

17. *New York Times*, 3 August 2001, p. A21; *New York Times*, 21 November 2001, p. A12.

第三章　媒体能帮忙吗?

科学是一部长电影，新闻媒体则通常拍快照。

——约翰·施瓦茨 (John Schwartz) [1]

公众理解科学真的有必要吗？为什么不让科学家做自己的工作，而其他人干好自己的事呢？不幸的是，在现代社会，那只是一条行不通的路线。不管我们是否认识到，科学已经成为我们生活的重要组成部分，我们不能将它当作好玩的附属品视而不见。与以前相比，经济、国防、环境和健康更需要依靠科学进步。信息技术在经济生产中的突出角色、弹道导弹防卫保护的可行性、人类燃烧化石燃料引起的气候改变、新的人类基因组图完成的含义，所有这些都提醒我们不能离开科学，即使我们想这样做。然而对于所有的科学与我们日常生活的明显相关性，许多人仍然没有准备在科学基础之外了解很多。

就像我在前一章讨论的，如果我们的学校不能发展对科学的意识

和评价，那么我们可以构想第二道防线来抵制科学文盲：科学家同大众传媒紧密合作来告知和教育公众。当科学问题的理解或误解出现时，我们能求助于电视、电台、报纸、杂志和因特网来澄清和深入了解吗？有了10亿美元的财政预算、有才能的工作人员和有时同实践科学家紧密的工作关系，人们理解科学的潜在性似乎很高。科学家和新闻记者通常都接受过良好教育，也具有相似的智力基础：好奇、怀疑和将不完整、有时是不准确的信息拼凑成一个故事的能力。的确，被学校忽视的科学教育，后来能够通过科学家和媒体的努力得到弥补。

对于科学家和媒体而言，这是个重大的责任，不幸的是他们都没有作好充分准备来承担这种责任。[2]科学家不善于交流，媒体内部充满竞争而没有耐心，这两个团体相互误解，在一定程度上互不信任，双方都缺乏教育公众理解科学的强烈责任感。再加上工作中许多复杂因素的影响，例如，像烟草行业和化石燃料工业这样有着特殊利益的群体，它们不想让某些科学问题得到澄清，也不想认出不确定性并对其评估。但是在媒体中，还有一些力量正有效地破坏着人们对科学的理解。广播和电视上的访谈节目主持人认为自己的首要角色是娱乐；科学很少有机会找到聊天的话题，而且科学通常很少能够在谈话中互动起来。并不是节目主持人有意将科学变得含糊，而是许多人不愿意把时间花费在理解科学的复杂性上。他们肯定不想让科学或其他任何被特别报道的话题，变得过于复杂或蒙上灰色的阴影。轻视性的简单化是一种逗乐的好方法，科学在一个能说会道的节目访谈主持人的解释下，往往很快被简化为一颗颗小石头。

这一想象的媒体-科学教育联盟也富含其他弱点。将这些有着共

同目标的潜在合作者联合起来，存在多方面的问题，其失败两个群体也都负有责任。俗话说，推卸责任应当受到谴责。首先让我们看看科学家设置了什么障碍。

科学家

新闻记者需要资料来源。如果科学家不愿意对他们讲科学的事情，他们就不能对此做报道。科学家喜欢他们的工作得到认可。难道他们能从媒体之外得到重要认可吗？不幸的是，至少从科学教育的观点来看，对这一问题的回答无疑是肯定的。

科学家寻求认可的主要形式是得到*同行的承认*。这主要通过以下方式来体现：他们的研究结果在经同行评议的科学杂志上发表；在同行评议的具有竞争力的提议的基础上获得研究基金；在同行评估基础上加薪和升职；对于极少数人来说，还意味着获得荣誉奖，如数学中的菲尔兹奖(Fields Medal)*，地球科学中的克拉福德奖(Crafoord Prize)**或者诺贝尔物理学、化学、医学、经济学奖。但是，通常对学院科学家的奖励体系并不重视非学术界的贡献。事实上，在学术生涯中存在一种潜在意识，即在非学术努力上花费时间会阻碍科学的进展。卡尔·萨根(Carl Sagan)是康奈尔大学天文学家，也是多产作家，他把许多科学带到普通大众王国，却从未当选国家科学院院士。当选国家科学院院士是美国科学家能获得同行认可的最高形式之一，

---

* 菲尔兹(1863～1932)，加拿大数学家。在数学方面的研究集中于代数函数。由他倡议的菲尔兹奖，是世界上最高的数学奖，同诺贝尔奖一样享有国际盛名。——译者

** 克拉福德(1908～1982)，瑞典商人。1980年，克拉福德捐献了300万瑞典克朗给基金会，也就是后来的克拉福德基金会。基金会的目的是促进诺贝尔奖之外的几门基础科学方面的研究工作。克拉福德奖从1982年开始，奖励类别包括数学、天文学、地球科学和生物科学(特别是和生态、进化有关的生物学)，此外还有多发性关节炎方面的研究工作。——译者

相当于英国的皇家学会会员。推断之所以对他拒绝认可，主要原因在于他的非凡成功无论是作为作家还是在电视上，都只是作为科学的普及者和解说者。许多成功的科学家，即便不是完全的不屑一顾，也对这类通过媒体的努力评价甚低。

尽管与萨根或古尔德(Stephen Jay Gould)* 相比显然是在很小的范围内，但我的确是在努力缩小这一差距。除了实际的教学工作和研究工作，我付出了很多的努力来为非科学读者描述科学、为非科学观众讲述科学，当我(或他人)的研究结果发表时，我会接受媒体采访，并帮助他们向公众传播研究的重要性。我已经从我们大学的副校长处收到愉快的信件，感谢我抽出时间同媒体合作(当然我的大学的名字也出现在文章或广播中)。我记得同克林顿总统科学与技术办公室主任、美国国家科学基金会前会长莱恩博士进行过交谈，探讨了我对本书的想法。他催促我立即动笔，并评论说帮助公众更好地理解科学是一件紧要的事情。我收到美国副总统的便条，感谢我为国会工作人员举办关于全球气候变化的讨论班。

自然，关于帮助公众理解复杂科学问题的重要性，写了许多，也讲了很多。然而，在承担着绝大多数科学研究的大学中工作的科学家，由于奖励体系的性质，丧失了参与媒体工作的信心。当同事们认真考虑成绩评估，将其作为年度薪金调整的一部分时，公共服务所占的分量很小。所有非学术认可再加几美金可以为你买到一杯上好咖啡，这种老掉牙的笑话绝非毫不相关。有分量的问题(文字上我使用这一术语)是：你发表了多少篇研究论文，你指导了多少研究生和

---

* 古尔德(1941～2002)，美国古生物学家，世界著名的进化论者、科学史学家和科学散文作家。主要著作有《自达尔文以来》、《熊猫的拇指》、《母鸡的牙与马的蹄》、《火烈鸟的微笑》、《为雷龙喝彩》、《八头小猪》和《鼎盛时期的恐龙》。——译者

博士后学者，你教了多少大学生，你获得了多少研究基金。更广泛的关于公众教育的问题，如本书部分章节所阐述的以及构成我的职业生涯一部分的更公众化的活动，在学校价值评估中很少得到较高评价。因此，遇到下列情况一点也不奇怪：如果科学家同媒体合作以便更广泛的非学术听众能够理解科学这件事情受到的实质性鼓励很小，那么他们很少再进行这方面的努力。

即使具有一个合适的奖励系统，训练科学家与媒体友好合作也不是一项简单的任务。也许首先遇到的困难是这一事实：通常科学家和新闻记者不是在同样的时间框架下工作。典型的是，科学不会产生诸如矿井坍塌、板球比赛、政治辩论或者芭蕾表演这样的日常事件来引起媒体的注意，科学家很少会面对按日计算的截稿期限。

也许将长期的科学努力制作成新闻事件的最恰时机，是当研究结果在会议上被提出或在专业杂志上被发表时。现在我们越来越习惯于新闻报道这样开始，“在今天发表于《新英格兰医学杂志》的一项研究中……”或者“本周在欧洲地球物理学会会议的一个报告中……”。但是，在每年发表的成千上万篇文章和论文中，媒体专业人员如何鉴别出其中真正有重要价值的呢？可以肯定，很少有记者会定期阅读《柳叶刀》、《美国医学会杂志》或者《地球物理研究杂志》，做这种事情对于大多数新闻记者来说，其痛苦无异于长期折磨致死。然而，近几年，大学、专业协会、科学杂志出版商作出很大努力，在他们知道和喜爱的科学家与向大众传播科学的媒体专业人员之间搭建桥梁。这是对绝大多数学院科学家和他们的专业协会所拥有的孤立主义传统的重大背离。

我自己所在的大学就很注意自己的公众形象，它有一个新闻和信息服务办公室，此办公室拥有一支能力非凡的工作队伍。他们未雨

绸缪，从科学部门负责人那里索取教职员参加的专业会议的日程表，然后与科学部门负责人坐在一起以便确认具有特殊重要性的陈述。接着就会对发表陈述的教职员进行采访，然后将准备好的新闻稿发送到媒体。有时这会导致对陈述之后的报告人做采访。

科学刊物本身也有给人以深刻印象、有效的新闻机构，促使媒体注意他们的出版物中出现的重大进展。两家享有盛誉的国际性科学刊物是出版于英国的《自然》杂志和由美国科学促进会出版的《科学》杂志。这些刊物在科学共同体内被广泛阅读，每周出版一次，特别推出15到20篇由世界各地科学家撰写的研究报告。在每周出版之前，《科学》杂志和《自然》杂志都提供各种论文的新闻稿、经过挑选的少量论文的编者按、使媒体能够与涉及该项研究的科学家进行交流以便获得其他观点的联络信息。最近我和同事在《自然》上发表了一篇关于“岩石温度如何证明地球变暖”的研究报告[3]之后，报纸、电视和电台几乎无休止地进行了将近一周的采访。

然而，如果没有科学家们的参与和合作，研究机构、科学协会及出版社再怎么努力也无法把科学带给大众。揭开科学神秘面纱的最终责任落到专业人员肩上。他们必须与报刊新闻记者、电视广播新闻记者进行有效的交流，而这些记者希望帮助科学家与公众一起分享科学殿堂里发生的一切。对于那些工作在大学的科学家来说，有效交流并非是不可逾越的障碍。他们面临的任务是始终把科学传播给教室里济济一堂的学生。成功的教师不会把自己藏在专门术语的障碍后面，而是学会组织和简化他们的资料，帮助学生既见树木又见森林。许多人已经认识到，将科学作为一种探究过程比作为成就的汇集向公众传播更为重要。他们知道所有的科学都是试探性和不确定的，正是这种不确定性刺激着创造性，推动科学不断向前发展。与

媒体合作需要遵循与学生交流相同的原则。科学家们必须警觉，并且对产生误解的可能性保持敏感，他们必须努力说清他们工作的意义以及不具有的意义。

记者们不仅需要科学家提供材料，还需要他们的深刻见解。对于记者来说，看杂志中带有某种神秘色彩标题的科学报告是一回事，而辨识出这篇报告的重要性(或不重要性)则是另一码事。然而，确实也有例外，有些研究只有活跃的研究者才能评估。事实上，这就是在文章刊登之前对其进行评估和筛选的同行评议系统的基础。相同的评估系统，通过征求活跃科学家对提交的寻求基金以便开展科学研究的提议的优点和弱点的意见，也指导着国家机构和州级机构的研究基金分配。我也相信大学层次的教学是通过积极地从事研究来提高的。没有人能比一位研究人员更能认识到大量实验数据的粗略性或者确认一个理论的薄弱环节。无论多么贪婪的读者，无论谁对科学著作多么能综合，也无法与那些工作在科学第一线的科学家相比。

记者们是忙于面对截稿日期的人，他们欣赏能切中要点的科学家。他们既不渴望听到也不擅长整理科学家创造的用于保护自己研究结果的无尽的限制条件。当记者们意识到结论有某种不确定性这样的事实时，他们也会感激科学家能把不确定性置于某种熟悉的背景中。霍廷斯基(Roberta Hotinski)是宾夕法尼亚州立大学地球科学专业的研究生，1999年，她整个夏天都在《美国新闻和世界报道》新闻工作室工作，该工作室是美国地球物理学联合会(American Geophysical Union)发起的，她的角色是担任科学与工程的大众传媒会员。她在描述自己的经历时，提供了以下建议："如果你用可能性或熟悉的比喻表达你的确定性水平，报道者将有一些具体的东西可以强调。

例如，你描述对全球变暖的确定性，可与你的信念作对比：a)太阳明天将升起；b)你的小孩将要上大学或c)你将要赢得彩票。"[4]科学家一般通过确定定量的范围的方式描述他们结果的不确定性，此量的范围被称为“误差棒”，他们的结果位于其中。把误差棒翻译成记者能够理解的普通语言，以便研究结果能够被理解，还要走很长的一段路。

媒体

阻碍科学和媒体紧密合作的所有障碍并不都来自科学阵营。让我们了解以下由媒体设置的几个障碍吧。就像施瓦茨在本章开头引论中评述的那样，媒体，拥有快速照相机，通常没有时间或耐心拍一部长长的电影。即使当一部长电影正在播放时，媒体也会因最近发生的事件而转移注意力，他们可能甚至都没有意识到自己是在剧院里。而且，由于全球气候变化这样的话题已经讨论了几十年，媒体认为此类话题缺乏新鲜感或流行性。但是，与气候相关的显著效应不断发生，从而可以推知气候变化也在持续发生，所有这些又把人们的注意力引向正在缓慢发生的变化。在议院或国会的另一个辩论是陈旧的，但是尺寸如苏格兰大小的冰山与南极冰层的突然分离却是新闻。一份来自一家国际石油公司要求为气候变化作更多研究的新闻稿让人厌倦，但是“发现从北冰洋到北极的路径上没有冰”的一艘破冰船又把媒体的注意力拉回到气候变化的另一个方面。而且当一个关于乞力马扎罗山*的流动冰的调查表明海明威（Ernest

---

* 乞力马扎罗山位于坦桑尼亚东北部，是非洲的最高峰，被誉为“非洲屋脊”、“非洲大陆之王”。海拔5895米，在茫茫无边的大平原上拔地而起，显得格外雄伟。山麓地处赤道，气温酷热，最高可达59 ℃，但峰顶温度在－34 ℃，终年积雪，寒风怒号。这奇异的自然现象使它又有“赤道雪峰”之称。——译者

Hemingway)笔下不朽的雪* 也许会在下个 15 年消失时，新闻室就有人记录下来。

很多串在一起的快照对于公众来说就像是一部电影。 通过重复，最终这些大尺度的环境概念能够进入公众的意识之中。 1987 年，当《蒙特利尔议定书》(一个规定逐渐停止制造和使用破坏臭氧的化学品的国际性协议)初具规模时，在国际会议之前有一个 10 年或更长时间的"快照"可见性是非常有意义的。 本尼迪克(Richard Elliot Benedick)写道：

> ……知识和公众观点的力量对于蒙特利尔的成功是个难以应付的因素。 见多识广的公众是推动政府的政治意愿和削弱工业界捍卫化学品的决心的先决条件。 科学家的发现不得不被了解和传播……媒体，特别是报刊和电视，在把这一主题带到公众面前和由此激励政治利益方面发挥着非常重要的作用。[5]

就像我已经提到的，许多小学和初中，实际上还有一些更高层次的学习机构，在科学教育和一般计算能力方面的不完备性现在已经影响了好几代学生。 大多数从业记者也包含在其中。 大多数媒体记者受到的科学教育与他们的小学、中学和大学同学几乎没有什么不同，这也导致他们随后在向公众传达科学的意义和重要性方面遇到困难。

---

* 海明威有篇名为《乞力马扎罗的雪》的短篇小说。 一开头寥寥几笔便勾勒出一种神奇的意境："乞力马扎罗是一座海拔 19 710 英尺的常年积雪的高山，据说它是非洲最高的一座山。 马基人称西高峰为'鄂阿奇-鄂阿伊'，意为上帝的庙殿。 在西高峰的近旁，有一具已经风干冻僵的豹子尸体。 豹子到这样高寒的地方来寻找什么，没有人作过解释。" 乞力马扎罗山紧靠赤道之南，它的诱惑力很大一部分来自赤道与雪山的强烈反差。 2001 年，根据《环球时报》报道，科学家发现：乞力马扎罗的雪正在融化，并可能在 15 年内完全消失。 ——译者

而且当科学家自己提供的一些东西缺乏一致的解释时，当不确定性被表达时，媒体报道者通常不能评估这些不同的观点，不能帮助公众理解和适应不确定性。

其他的因素与教育之不足也不是没有关系，包括将科学视为无趣和很少有新闻价值的过低评价。这一态度显然是由于工作人员缺乏约束、科学报道的时间和出版者、制造者、编辑的分析引起的。这一态度受以下观点的支持：科学在日常基础上受时间和空间的限制，实际上不具有竞争性。另外，不幸的是，一些媒体的“看门人”感到公众教育不是媒体的责任；他们认为他们的责任是报道和记录每天的事件，而不是解释那些事件和将事件置于相关背景之中。更有甚者，他们仅仅希望取悦其读者。

对于一个以熟练使用语言为荣的职业，媒体在评估特殊利益群体使用的语言时经常粗心大意，这种语言影响了描述问题的方法。这些巧妙歪曲主题的代码字常常被无意地并入文章。将大气的二氧化碳和甲烷描述为“所谓的温室气体”，产生了一种不该产生的印象，大气科学家绝对不支持这种印象。当记者使用诸如“温室理论”的词汇时，他们产生了一种印象：大气温室效应也许只是一个概念，根本没有现实基础。没有什么更加远离真理；纵观地球的历史，地球表面由于温室效应而变暖。合适的辩论不是我们这颗星球是否具有温室效应，而是温室效应因为人类活动正在改变多少。重复使用“健全科学”、“垃圾科学”和“创世科学”这些词语，而不仔细留意是谁在使用这些语言以及他们可能的动机是什么，对于公众来说是极为有害的。

媒体经常认为争论比科学本身更有趣。科学辩论差不多就像一场体育竞赛，由于具有竞争性肯定会产生一个获胜者。但是不像每

天都可以产生一个非常重要的报道的体育运动，科学主题没有大规模媒体队伍和深度分析的益处。大型报纸也许拥有一个体育记者团队，这些记者都是高尔夫球、网球、板球、足球、篮球和棒球方面的专家。许多星期以来，我们也许了解到有关训练课程、训练战略和世界杯足球赛球队的群体心理，或者一名环法自行车赛选手深刻的内心思想的许多细节。但是科学评论经常如此报道“他说，她说”，很少出现提供给读者或观众以便帮助他们理解奥妙的深刻见解。因为新闻记者希望避免成为一边倒的拥护者，他们承认自己不能充分地分析形势，因此经常感到必须拿出相等的时间给相互对立的观点，不考虑支持科学争论的任何一方。结果常常导致将注意力不必要地投向并由此相信一些轻率的宣言和不重要的争论。

媒体新闻走入歧途

之前：“断层线的威胁一触即发”[6]

之后：“媒体在新马德里地震的错误报道让人吃惊”[7]

媒体新闻走入歧途方面一个具有教育意义的例子是，预言 1990 年 12 月 3 日发生的一次大地震将会摧毁美国中部大陆，或中加利福尼亚，或东京，或其他某个地方所带来的威胁。这一预言是由布朗宁(Iben Browning)博士发布的，他是一名商业顾问，其最高学位是在生物学方面获得的，而不是在地质学或地震学方面。在美国，特别是中部的州，媒体报道这一预言时并没有做任何重大努力来评估它的可信度或发生的可能性，结果酿成了大事件，导致学校关闭、商铺停业、预防突发训练、保险欺诈，以及在地面零点利用几乎像狂欢节一样的氛围举办非同寻常的企业活动，给当地造成了混乱。

布朗宁关于地震的预言基于这一思想：1990 年 12 月 3 日，太阳、月亮和地球将排成一线，结果太阳和月亮的万有引力将达到非同寻常的水平。这些额外的作用力将加到地球地震带不断累积的结构压力上，从而引发一次大地震。这是一种“压垮骆驼背的最后一根稻草”的概念。虽然布朗宁让人们注意的行星排成一线的现象的确发生了，然而作出这一预言只需要读一本历书，不需要更多专门的科学技能。布朗宁有什么能够证明预言将要发生地震的地方已经达到断裂点，正等待最后一根稻草致命一击的证据吗？绝对没有。显然他只是选择了历史上曾经发生过大地震的地区，宣布这些地区将面临又一次的冲击。

幸运的是，12 月 3 日来了，又安然无恙地走了，在预言中的密苏里东南的震中(在世界任何其他地方也未发生大地震)没有发生一次能被察觉的震感。震中地带马戏团的帐篷收起来了，媒体又把它们的注意力转向随后发生的其他更具“新闻价值”的事件。原本能够为公众提供更好的服务吗？绝对可以。记者有许多机会更深入地探测新闻报道中的问题和实际判断的科学基础，但是通常都没有做到。

媒体可能会做什么？布朗宁地震预言中的“可信性”，显然起源于他于 1989 年 10 月 10 日对旧金山制造业经理的讲演，在讲演中他宣称一周之内世界上某个地方将会发生一次大地震。果不其然，那周就在加利福尼亚发生了洛马普列塔地震，导致 67 人死亡，给旧金山造成巨大破坏。但是这一“预言”真的就这么成功吗？仅仅作了一点调查，记者就可非常容易地了解到在全球范围内平均每年大约发生 120 次大地震(震级 6 级以上)，大约每 3 天就发生一次地震(为了简便起见，假设地震在时间上均匀分布)。因此布朗宁可以高枕无

忧地对时间做出合理猜测。一个初出茅庐的记者同样能够有这样的先见之明。

除了指定地震发生的“时间”，地震学家需要对以下因素作出可靠预测：地点、震级以及那种规模的地震发生在那个地点的概率估计。就布朗宁指出的地点而言，世界上某处确实是一个安全带，预报“大的”地震为解释地震活动的大小留下了回旋的空间。如果目标足够大，就很有可能实现。至于对概率估计的需求，可以用一个例子很好地说明：旧金山湾地区将在某天经历一次 2 级地震，即使果真像预言那样地发生了，也不会把它当成有任何重大意义的预言。为什么呢？因为在那一地区每年要发生几百次 2 级地震，所以在任何给定的一天发生地震的概率都是非常高的。如果布朗宁在他对商业经理的讲话中预言在即将到来的一星期内中加利福尼亚将发生一次 7 级地震，这就与声称世界上某个地方将会发生一次“大”地震完全不同了，接着洛马普列塔地震的发生将会显著提升他作为预言家的地位。标枪直接击中一个小的目标要比击在一堵大墙上某处给人留下的印象深刻得多。

科学共同体一致认为：布朗宁的预言不具有科学价值。那么为什么媒体对这一事实大打折扣？美国地质调查局国家地震预测评估委员会宣布布朗宁的预言无可信度，附近大学大多数专业地震学家也是这么认为。然而这里有一个例外，一位来自东南密苏里州立大学的地震学家，对布朗宁的预言给了相当大的公开支持。对于媒体来说，这一孤独的支持者很明显地对坚定坐在跷跷板另一端的大众科学观点提供了一种应有的抗衡。如果媒体对这位持支持立场的地震学家的背景做深入调查，那么他们会发现此人对地震预言的业务并不陌生；他已经同一个通灵人合作预言了 1974 年北卡罗来纳州的一次地

震。如果媒体深入研究布朗宁博士自己的背景，他们也会发现他把纳粹党人的出现归因于潮汐力！无论发生什么情况，记者都要对背景和资源的可信性进行检查。

布朗宁曾经预测1990年12月3日，行星排成一线会对那个地区施加一个强大的万有引力。媒体原本也可发现在这一预言背后是陈旧的和不成功的科学。由于对太阳、月亮和行星的位置能够进行可靠的预测，因此这总是其他预测主题的极具吸引力的组成部分。在有关过程的物理性质能够被很好理解的自然系统中，诸如海洋潮汐沿着海岸线以及海港的潮起和潮落这样的预言取得了显著的成功。但是在关于灾害的预测业务中，由于自然系统复杂和难以理解，记录显示没有成功。

有一本在1974年出版时受到广泛关注的书[8]的作者说，行星排成一线将在1982年发生，这种结构每隔179年才发生一次，它将在南加利福尼亚引发一场灾难性的地震。这种引发地震的物理过程比布朗宁在他的1990年预言中所包含的物理过程更复杂，但宣传是同样的：

> 地球物理学家报告［圣安德列亚斯断层］*早就成形了……一触即发。我们感到没有什么疑问的是：在罕见的行星排成一线之后，行星和太阳的影响……将会触发。特别地，我们认为洛杉矶地区将要遭受本世纪大型人口中心经历的最严重的地震。[9]

---

* 圣安德列亚斯断层：断层实质上为地壳上的裂缝，它们的长度有时候会长达数百公里。当覆盖着地球表面的板块(tectonic plates)移动时，断层所在的位置，常常也是大地震发生的地点。圣安德列亚斯断层是美国加州著名的断层，它也是地球表面最长和最活跃的断层之一。圣安德列亚斯断层的深度有15公里，存在的时间已经超过两千万年。巨大的太平洋板块(Pacific Plate)沿着圣安德列亚斯断层，相对于北美板块(North American Plate)向北漂移，平均每年移动数厘米，以这种移动速率经过数百万年后，地球表面的陆块分布和现在比起来，就会有很大的不同。——译者

不用说，1982 年南加州的圣安德列亚斯断层在 1982 年并没有剧烈震动，洛杉矶也仍然屹立。但是这一事件在布朗宁给出他的 1990 年的预言时，很大程度上（彻底地？）已被媒体忘记了。新闻编辑室文件箱内标签为“地震预测”的淡黄褐色文件夹，记者本可用来查阅背景和历史观点，却明显是空的。我没有料到在布朗宁的媒体疯狂时期对木星的“无效果”仅提到一次。有另外一本具有类似主题的书，[10]预测 2000 年 5 月 5 日将有一次全球性灾难，尽管这时媒体新闻[11]限制并避免宣传，又一次显示出对这种特别灾难主题的频繁再现没有显著意识。

为什么媒体不去揭开布朗宁预言的假面具，说它的荒谬可笑呢？为什么他们反而将他描绘成一个偶然发现“预言地震的超常简单策略”的非正统的、未被欣赏的天才？的确，要回答这些问题并不很简单，但是一个方面也许是对专家队伍和传统学问根深蒂固的不信任。在许多背景下，这种态度能够被描述为“反精英主义”、“怀疑主义”、“敌对主义”或者“对弱者的支持”。如果所有大人物都说某件事不会发生，那么当此事发生时，看着他们丢脸难道不可笑吗？当没有成为团体会员的某人不顾轻蔑的反对而取得成功时，难道我们不可以开怀大笑吗？当所有专家都告诉我们某件事非常复杂，让一些不被承认的外行人表明这件事情其实非常简单难道不好吗？复杂性（complexity）是理解的真正障碍，它引起人们寻求更简单的答案，而不管这些答案是多么的不充分或者具有多么明显的错误。

这种态度以一种有趣的方式展现在地震预测的另一个例子中。20 世纪 70 年代早期，在美国矿物局工作的一名科学家，在实验室压力测试中有关岩石如何破碎的研究基础上，建立了一套地震预测理论。接着他把他的理论应用到现实（实验室外）的世界，1976 年他预

测一场20世纪最大的地震将于1981年6月28日在秘鲁海岸发生。[12] 很自然地，秘鲁人对这一即将到来的灾难所进行的预测非常关注，并对其价值寻求评估。因为这一预言由一名美国政府的科学家发布，所以起初秘鲁通过外交途径同美国进行了协商。美国政府转而让美国地质勘探局——掌管评估地震灾害的联邦机构，评估这一预测的理论和方法论。美国地质勘探局评估组做出结论，称那位美国矿物局科学家的方法不可信，不应相信他的预测。特别是，他们下结论说秘鲁在历史上曾经发生过严重的地震，该国的地震灾害绝不会在美国矿物局科学家概念的基础上加剧了。

事情到此并没有画上句号。一些人开始是私下里后来就公开质疑美国地质勘探局否定美国矿物局的方法是否是出于嫉妒。众所周知，美国地质勘探局在地震预测研究方面已经投入了相当大的精力，却收效甚微。如果美国矿物局单个的研究者发现了地震预测的秘密，然而整个美国地质勘探局却没能做到，那不是非常尴尬吗？这不就是一个小伙子主要利用业余时间成功地阐明了一个复杂的过程，而所谓的专家仍然在一条又一条黑漆漆的小巷里摸索的实例吗？一个不被赏识的天才不是让专家显得很无能吗？这一事件具有后来的布朗宁传奇故事的所有特点，包括将要发生的预测事件的明显失败。那一天到来了又过去了，在秘鲁没有感受到多少地震波。正像地震专家预测的那样，美国矿物局的预测结果毫无意义。

正确是非常困难的

就像让科学家成为媒体友好的同伴遇到的困难一样，让记者成为科学的成熟评论员同样是困难的。我们希望记者不只是文字报告者，他们应当能够提供一些见解和观点。记者最出色的状态是必须

具有怀疑性，而且能够搜索出弱点和矛盾。对科学的理解和熟悉达到一定水平的那些人，做了以下典型评论：达到这么舒服的高度是一个多么艰难的过程。《纽约时报》的布朗（Malcolm Browne）评论说："一个科学作家必须是永久的学生。……意识到一个精妙实验的价值需要有准备的头脑。"[13]克朗凯特（Walter Cronkite）为国家电视网报道美国的空间计划，他通过"很长时间的学习"战胜他对科学的紧张情绪。[14]但并不是很多记者都有增强科学背景和洞察力的机会。更为典型的例子是一个新的年轻科学记者的经历，她刚开始工作时感受到巨大的压力。几星期内她不得不报道关于大脑的疼痛中枢、低频电磁波辐射对人类健康的影响、测定沉积岩年代的新技术、阿耳茨海默病*的潜在疫苗、纤维光缆的使用和全球变暖的预测[15]等内容。

幸亏，也有一些引人注目的例外：伦敦的《泰晤士报》和英国的《卫报》、加拿大的《全球邮报》、美国的《纽约时报》都提供了超出一般水准的科学新闻。当然，2001 年人类基因组图谱的完成几乎在世界各地都被当作头条新闻来报道，但是《纽约时报》却又用了满满 10 版的篇幅来讨论这一杰出科学成就的医学意义和伦理意义。过了大概一个月的时间，《纽约时报》在头版突出刊登了几个不同的科学故事：火星干冷的侵蚀特征的发现，暗示地质上较近时期的水是从地下渗出的；对暗示鸟类有着一种不同进化起源的一些羽毛化石需要重新解释；中子家族中发现了一种难以捕获的粒子；基因检测的社会后果。另外，每个星期二，《纽约时报》都会提供一个完整的版面——

---

* 阿耳茨海默病（Alzheimer's disease），俗名帕金森综合征。是一种进行性神经退化性疾病，主要症状为记忆力减退，随着智能慢慢衰退，最后甚至忘记吃饭、忘记走路，其临床表现多为运动功能的缺损，包括颤动、僵直、运动迟缓、运动技能减退、姿态异常等。——译者

科学时报版，该版的文章由15名科学作家共同撰写。

但这就是要庆祝的理由吗？ 在英国只有一小部分的报纸读者阅读伦敦《泰晤士报》或《卫报》，在美国阅读《纽约时报》的人可能就更少了。 而且所有这些报纸都用比建立在每周基础上的科学报道大得多的空间报道建立在日常基础上的财经新闻、娱乐、艺术、体育。 2000年7月21日，《纽约时报》有5篇重要的科学或与科学相关的文章，对于科学可见性来说是非常好的一天。 然而，就在同一天，有38版是关于人文和娱乐，6版关于体育。 而且我们必须记住，绝大多数的媒体分部甚至没有雇用一名科学记者；他们只是从通讯社简单地拷贝故事，然后不加选择地随意切割以填充能够获得的空间或时间。 大多数地方电视频道都有一位“气象主持人”，此人也许会有一个气象学学位。 无论何时当科学闯入日常程序时，常常会让气象员或女人来处理它。

我有点悲哀地得出结论：科学教育结构的缺陷不会因为获得媒体的援助而奇迹般消失。 无论是科学家还是记者，对其不作出足够努力都有着自己的原因。 指望实践科学家和媒体在科学教育停止的地方继续努力，我想恐怕是太奢华的一个梦。 那里可能有潜在性，但是要想实现，就像是海市蜃楼一般虚幻。 科学教育和媒体对科学的关注会得到改善，我对此满怀希望，但要彻底转变，路还很漫长。

现在是离开东方帐篷，进入到科学不确定性花园的时候了。 就像前边提到的，花园中有许多区域和小块土地显示了不确定性的多面性特征。 然而，许多花卉陈设区和未被注意的自然区域，灌木丛和巧妙的迷宫，它们通过在各自不确定性的区域突出的图像共同构成了一幅镶嵌画。 它们一起展示了不确定性如何产生、如何刺激创造

性、为什么类似于一个半满而不是半空的杯子。接下来的一章，将探索在理解复杂现象时我们的直觉引导常常是不够的，由此允许不确定性覆盖着我们日常经验之外的不熟悉领域。

注　释：

1. *Washington Post*, 21 February 1999.

2. Jim Hartz and Rick Chappell, *Worlds Apart: How the Distance between Science and Journalism Threatens America's Future*, First Amendment Center, Vanderbilt University, 1998。其中对科学家和媒体从业人员的关系进行了透彻的讨论。

3. Huang, S., Pollack, H. N., and Shen, P.-Y., Temperature trends over the past five centuries reconstructed from borehole temperatures. *Nature* vol. 403, pp. 756 – 758, 2000.

4. Roberta Hotinski, *EOS Transactions of the American Geophysical Union*, 16 November 1999.

5. *Ozone Diplomacy*, Harvard University Press, Cambridge, 1991, p. 5.

6. *USA Today*, 28 November 1990.

7. *St. Louis Post-Dispatch*, 8 December 1990.

8. John Gribbin and Stephen Plagemann, *The Jupiter Effect: The Planets as Triggers of Devastating Earthquakes*, Walker and Company, New York, 136 pp., 1974.

9. 引自 *The Jupiter Effect* 的前言。

10. Richard Noone, *5/5/2000: Ice, The Ultimate Disaster*, Harmony Books, 1986, 1997.

11. *New York Times*, 7 May 2000, p. 26.

12. 要全面分析这一事件，参见 Richard Stuart Olson, *The Politics of Earthquake Prediction*, Princeton University Press, Princeton, NJ, 1989, 187 pp。

13. *New York Times*, 27 February 2000.

14. Jim Hartz and Rick Chappell, *Worlds Apart: How the Distance between Science and Journalism Threatens America's Future*, First Amendment Center, Vanderbilt University, 1998.

15. Roberta Hotinski, *EOS Transactions of the American Geophysical Union*, 16 November 1999.

第四章　不熟悉导致不确定

经历就像一名严厉的老师，因为她先进行测试，然后再讲课。

——弗农·桑德斯·劳(Vernon Sanders Law)

当我们在生活中经历一些事情时，我们会熟悉它们，也许会理解它们，进而接受它们，并把它们当作生活中很自然的一部分。但当我们遇到先前从未经历过又不熟悉的某事物时，很自然我们会怀疑它。而且如果我们面对的是一个抽象概念的话，而这又完全超出我们的经验，那么怀疑或甚至不相信都是一个自然的反应。

在这种情况下，不确定性与不熟悉紧密相关。对于我们所不熟悉的事物，我们对其也不确定。而且科学中的很多东西对于许多人来说是不熟悉的领域。尽管爱因斯坦(Albert Einstein)不这样认为，但科学的确并不仅仅只是“常识”。如果科学真的是常识的话，稍有一点常识的人也不会对科学感到迷惑。科学需要一定的抽象，需

要把观察数据放到一定的背景或框架中去。当框架是一个人目前的环境，这就很容易熟悉和理解。但当空间框架像在粒子物理中那样小或者像在天文学中那样大时，就只有那些有心人才能探索这个不熟悉、不确定的领域。同样的，它们运行过程的时间尺度与日常人们经历的时间尺度有很大不同。瞬间化学反应的完成或者非常缓慢的地质变化都需要智力的延伸。

经历能够将不熟悉的东西转化成熟悉的东西。早期人类组织他们经历所处的概念框架不是抽象的。生存要求他们敏锐地观察周围的环境。对于他们来说，捕获猎物以及避免被猎捕是生死攸关的事情。充分认识周围的环境是生活的必需。当早期农业刚刚出现时，那些整年能够察觉当地的降雨模式、干湿季节的交替、气温的季节性变化的人，就更能成功地生产食物。但是早期的人类关心过全球温度长达一世纪的变化吗？尤其是当温度每隔 24 小时有 20 ℉或 30 ℉的改变或者季节性变化更多时，他们能够记录、甚至关注今年的年均温度比去年高了十分之一吗？或者他们能够想象出他们对土地的耕耘将对地球的另一边产生影响或改变全球大气的化学性质吗？

当家里发生大事的时候，人们就很难关注全球范围内增长很小的变化。当最大的威胁是地方性的并日益增大时，处理眼前事情的策略非常有效。例如，在现代社会，当人们被要求考虑“在行星范围内且时间尺度超出了常见的政治术语的界限甚至几代人寿命的范围”的全球气候变化这一概念时，人们往往会犹豫甚至产生怀疑，因为这超出了由日常生活得出的日常经验的范围。因此我们要慎重，这是一种小心翼翼进入到不熟悉领域的自然趋向。不确定性伴随不熟悉。

引起我们的注意

我们可以调节感官来察觉快速的变化。我们开车时，对汽车的喇叭声、警报器的尖叫声、车前方突然亮起的刹车灯、冲到街上的狗，都很警觉。所有这些变化都是在我们称之为“正常”的背景下进行登记的。喇叭声或警报器声是发生在日常持续不断的交通噪音的背景之下的；刹车灯在先前没有照明的地方会闪光；狗也是作为一个不同尺寸的物体在横穿马路的方向上快速移动突然出现的。每一个对正常背景的干扰都引起了我们的注意，警告我们可能有危险。（相反，当我们想入睡，并想放松警惕时，我们可以借助于听音乐或电台、电视上无须留神的谈话，或者是聆听特殊设计的波涛拍打海岸或风吹过叶子的声音等背景音。）

当高速公路上发生车祸时，随后的车流会立即减缓并开始阻塞。同样地，当我们开车到达两个车道合为一个车道的交叉口，并且这个车道只允许一辆车通过时，我们注意到了减速。当修路工人在施工和交通阻塞发生时，我们注意到了一天中工作往返的路程要比往常多花费半小时的时间。但我们是否注意到我们日常往返的路程因为城市的扩张而非常缓慢地延伸？我们是否注意到今天在路上花的时间与高速公路上车辆少得多的10年前相比要多出30分钟？答案也许是否定的，因为这种变化进行得很慢。由于微量延误的积累，日益增长的交通的放缓效应出现了，但却因为建立在日常基础上通常未引起注意。我们的感官还没有进化到让我们注意一些微量增加的变化，而这些变化又是经历了很长的时间间隔。只有通过我们个人的记忆，或者通过我们所谓的历史记录的集体记忆，我们才能注意到日常生活背景下的缓慢变化。当高速公路工程师发表报告说，15年前设想并建造的每天接纳10万辆车的州际高速公路系统中的一段，现

如今承载的数量已经超出原设计的两倍时，我们才想起这些变化。呵！原来这就是现在我每天上班时花在路上的时间要比高速公路刚开通时多半小时的原因。[1]1999 年美国人平均花费在交通堵塞上的时间是 36 个小时，而在 1982 年这个数字只有 11 个小时。[2]人们之所以接受和适应这一路上花费时间的增加，部分是因为这是许多年来逐渐积累而成的。

气候是你期待的事情，天气是你得到的东西

基于同样理由，要察觉到气候的变化也是同样困难的。气候不同于天气。天气的变化是显著的——昨天是大风，今天是暴风雨，明天阳光灿烂。因为天气是每天变化的，我们能够察觉到它的变化，并乐于知晓每天天气预报中提供的未来几天天气的大概情况。

相比较而言，气候是“平均”天气的长期特征。我们使用大陆性、内地的、海洋的等术语来描述一个地区的气候。每一个术语都隐含地描述了年平均温度、季节性温度变化范围、年均降水量、地面有积雪覆盖的平均天数。当我们计划围绕冰岛进行一次夏季自行车旅行时，我们需要查阅相关的气候图册以便知道需要带什么东西。这样的图册将告诉我们那儿日常温度在 5 ℃到 25 ℃(41 ℉到 77 ℉)之间，任何一天都有 30%的降水概率。

这种类型的信息是气候性的，因为它描绘的是长期的平均情况。气候上的缓慢变化是很难察觉的，因为它并不是建立在具有显著性的日常基础上的。只有当日平均温度被观察到有上升的趋势或几十年来年降水量有下滑的趋势时，那些保存着日常气象观测历史记录的气候学家，才提醒我们注意：我们正在经历一个气候变化。

但也有一种倾向是把短期内偏离正常看作是长期的趋势。北美

中部大陆的许多居民会把1988年的夏天回忆成是苦难的一个预演。那年夏天气温达到了历史记录并又保持了很长的时间。我住在密歇根，位于北美五大湖区*当中。那里的居民忍受了持续40多天超过32 ℃(90 ℉)和10天超过38 ℃(100 ℉)的高温。草晒死了，树枯了，庄稼枯萎了。报纸称全球变暖已经到来，这是对人类的报复。相比较而言，在2000年的夏天，没有一天的温度高达32 ℃，雨水也很充足。很显然，无论是1988的夏天还是2000年的夏天，都不能决定长达一个世纪的气候趋势，气温超出其年平均温度的任何一个地方都不能决定全球的平均温度。至少在确定长期变化趋势的背景下，我们必须小心，不要对短期现象给予过分关注。

对于个人而言，要察觉出年平均温度上的长期缓慢变化是非常困难的，部分原因在于把这些变化同昼夜之间，或冬夏之间发生的温度变化相比，这种变化是很小的。从生理学或心理学上讲，我们都无法注意到加在巨大迅速变化上的那些微小缓慢的变化。“不要拿小事来烦我”这句熟悉的格言表达了人们有以下自然倾向：关注现在正发生的大事，而不是长时间发生的微小变化，即使这些不断增加的微小变化也许最终具有一个大的影响。其实，即便是“大海中的一滴水”，如果水滴积累足够多，也可以让海平面上升到可以觉察的程度。

另外，生活在现代技术世界通常意味着，就温度而言，我们把自己隔离在自然界发生的变化之外了。这种隔离来自于这一事实：我们中的许多人生活和工作在经“气候控制”的环境中，即大楼里。

* 是指在北美中部美国和加拿大之间的五个淡水湖，包括苏必略湖、休伦湖、伊利湖、安大略湖和密歇根湖。——译者

在我们的温度隔离中，我们设定和调节机器让室内保持一个狭窄的温度范围，其结果就是我们接触不到自然界的变化。可以想象人们能够意识到年平均温度的长期变化趋势的唯一方法，是通过测定控制我们室内气候的机器消耗的能量多少的变化来实现。

超出范围

如果说我们意识到发生在我们周围的长期变化是困难的，那么意识到地球上其他地方所发生的事情同样也是困难的。事实上，人类完全可以想象我们在某地区经历的事情也在世界上其他地方发生。如果我们在密歇根有一个很少降雪的暖冬，那么很容易想到其他任何地方也有一个类似的暖冬。然而，可以告诉一下生活于西伯利亚的人们以下事情，他们在同样的冬天要经历寒冬和创记录的降雪。

我们作为个人所经历的是一个全球性的过程在局部的表现。地球上区域气候变化的第一级模式是赤道地区气候变暖，极地地区变冷，这通常反映出地球表面每平方米接收到的太阳能的数量变化。然而主要由地球旋转轴的倾斜引起的半球的季节变化和通过大气环流和洋流进行的地球热量的再分布也改变了这一受太阳影响的基本模式。大气环流和洋流进行的地球热量的再分布尤其受大陆地形所影响，它们限制和引导海水的流动方向。

随着这样的地区性变化的发生，一些地方变得比平均冷些，而另一些地方比平均暖和些，因此全球平均温度的概念就变得有点抽象（在下一章中我将描述全球平均温度是如何通过许多地区观察数据累积获得的）。以下情况是非常可能的，地球上没有一个地方每天的温度实际上都是全球平均温度。但是在一个有 30 个学生的教室里，人们可以计算出这些学生的平均身高和体重，即使这个班上没有一个人

的身高和体重与平均值一样，上述情况并不比这更为抽象。

尽管我们经历的天气是全球气象系统的地区表现，尽管环境学家一再劝告，但是我们仍然不习惯进行全球化思考。然而我们的视野范围依靠环绕地球的卫星已经大大扩展了，能够让我们全球性“察看”。从离我们22 000英里的卫星上可以观察到整个大陆的天气状况，现在我们每天也可以通过电视或者互联网来了解。我们能够看到形成于大西洋的飓风向西不断推进，一直到美洲东海岸的某处登陆。同样的卫星技术使我们能够容易地、快速地、清晰地观察到世界的每个地方。下述情形都是过去的事情了：一个人要想从非洲或南美向欧洲或北美打电话，不得不提前一两天预约，而且当最后接通时，声音效果听起来很差，比我小时候玩的一种“用罐子听声音”的游戏声音效果好不了多少。手机、电子邮件、经由卫星传送的电视信号为全球所有地方带来了即时通讯。巴伦支海的一个水下灾难，南极的一种疾病，哥伦比亚灌木丛林中的游击战争，或者南太平洋一座新喷发的火山将其火山锥喷出海平面，所有这些现象现在在地球上任一角落都能立即被观察到。

远程通讯革命引起了全球的经济变化，而这又冲击了我们地理上的乡土概念。全球化现在是一个用于描述国际商业和贸易关系的常见词。跨国公司、政府和工人都意识到世界上某一个地区所发生的事情会影响到世界所有地区。东南亚的货币不稳定会在全世界的交易所引起金融风暴；伦敦和香港的债券价格的微小差异会导致瞬间内大量资金的电子转移。有关国际贸易团体会议以及由其引发的抗议浪潮，成为头版要闻，不再被深藏在报纸的商业版块内。

然而颇具讽刺意味的是，全球化也有微妙的方面，使得我们很少意识到经济和气候的地区性变化。在地球上人类的史前早期，当农

业成为人类食物供应的重要组成部分时，长期的干旱会迫使人类迁移到降水量多的新地区发展农业。那些迁移村落的成员一定会意识到地区气候的变化和气候变化对他们的生活造成的压力。今天，在全球一体化的经济中，一个地区的产量减少能够通过其他地区的增长来弥补。就消费者而言，无论产量是减少还是增加，都离消费者很远，反正供应看起来是稳定的，所以他们对气候变化状况漠不关心。对于典型的消费者来说，橘子是来自佛罗里达还是以色列并不要紧，只要它们定期出现在超市的货架上就可以了。

全球化并不仅仅意味着通过远程通信把世界连接在一起。它有时也被看作是“边界的终结”。在欧洲的大部分地区，很少需要签证，许多国家的货币已经被一种单一欧洲货币——欧元代替了。它意味着在大陆内和大陆之间的人员和货物的不断流动，也导致了不受欢迎的“搭便车的旅行者”(hitch-hikers)的全球化：病毒的传播把疾病带到了以前从未出现过这种疾病的新地区。诸如斑马贻贝等外来物种已经在北美五大湖区找到了新家，这些来自欧洲的客人是通过货船船底的污水来到这里的。2001年蔓延于不列颠群岛的口蹄疫疾病的爆发夺走了很多农场动物的生命，它是从亚洲经过迂回路线传入的。

然而，全球化的环境方面与经济方面相比，引起的注意少多了。公众对环境问题的全球化所知甚少。公众很容易理解全球经济的观点，但是要让许多人思考全球性环境看起来要困难得多。为什么在公众的心里，对于全球气候的变化，或者全球的平流层臭氧损耗有这么多的不确定性？当然，这种非连接性，部分起因于为了既得利益，有意竭力传播对这些问题的怀疑和迷惑。然而对于怀疑论还有一个更深层次的原因，就是认为人类在巨大的自然力面前是软弱无力

的。人类只能眼睁睁看着他们的房屋在洪水或龙卷风中消失，他们的通讯和能源供应线路被一场冰暴毁坏，他们的高速公路因暴风雪而关闭，他们的海滨在一场飓风中消失，他们的村庄被山体滑坡埋葬，他们的城市被地震摧毁。一个人很难把他或她自己想象成大自然的弄潮儿。

当然，在某种程度上，有着上述想象的那个人是正确的。单独的他或她不是自然中的巨大力量。然而联合起来的话，情况就完全不同了。现在地球上的人口数量已经超过60亿，对环境具有巨大的影响。通过引入损耗臭氧的化学品和导致气候变化的温室气体，大气中的化学组成发生了深远的变化。海洋中的噪声污染，来自于轮船昼夜不停地穿梭、近海处石油钻井平台的钻探和抽取、设计用于让“全世界听到”的水下传导物发出的强烈信号，所有这些已经导致了一个令人震惊的情况：无论多远，没有哪个海域能逃脱人类活动所产生的声音。

人类作为一个整体正在全球环境中留下他们的印迹，其功效至少与人类通过远程通讯和经济在地球一体化中显示的那样大。然而，我们在承认“我们整体的自然力量在改变全球性环境”方面是滞后的。

以蜗牛的速度

如果一个人很难判断长达一个世纪的变化趋势的原因在于他的寿命通常小于一个世纪，那么判断更长时间的趋势就几乎是不可能的了。然而这里有更长期的力量在起作用，它们在地球气候中扮演一个重要角色，提供了一个非常慢的变化背景，在这个背景中，仅仅在一个世纪内发生的显著变化趋势看起来像是在稳定的交通嘈杂声中发出的刺耳警报声。长达一个世纪的变化趋势于是成了吸引我们注意的事件。

影响地球气候的长期因素来自地球围绕太阳运动的轨道形状的变化和地球相对太阳方位的变化。地球围绕太阳的运行轨道几乎是但不完全是一个圆，而且这点对于完美圆形的最细微的偏差就使轨道呈现轻微的椭圆形，其结果就是地球与太阳的距离在一年中并不非常固定。当地球离太阳近一点时，它就能获得较多的热量；当距离远一点时则少些。一年当中太阳热量的变化对于地球上出现我们称之为季节的现象起到一些作用（但形成季节性的更重要的因素是地球旋转轴的倾斜度）。地球轨道呈椭圆形的显著之处，至少根据气候领域的长期趋势来看，就是这个椭圆会随着时间的推移非常缓慢地发生变化。这个围绕太阳的轨道在形状上会发生波动，拉紧、放松、拉紧、放松，如此循环。一个完整周期大约起码要花 10 万年。在这个进程中，当轨道的拉紧在最大限度时，辐射热量的年度变化最大；当放松时，阳光强度的年度震荡就会减小，使得一年到头地球热量更加均匀分布。

当地球和太阳这种长达 100 000 年的运动发生时，地球本身也在按照自己的某种节奏运行着。地球每天旋转所围绕的旋转轴，今天正在偏离与地球环绕太阳的轨道平面相垂直的线。这一倾斜导致半球的季节变化；如果地球是垂直的，不存在倾斜，那么，“多伦多的冬天正是布宜诺斯艾利斯的夏天”这种现象将不复存在。但目前大约是 23.5°的倾斜角度，也是慢慢在 22°和 24°之间变化的，大约每 40 000 年完成一个周期。向 24°倾斜使季节对比更为明显，然而向 22°的倾斜则减少这一季节性对比。

在这行星运行进程中仍然还有第三种运动。旋转轴，每 40 000 年或上或下，同时也在作一个 25 000 年的脚尖旋转，类似于一个旋转陀螺的慢速摆动。其结果是引起了一年中季节的缓慢变化。尽管今

天北半球在6月、7月和8月是向太阳倾斜，北美洲人、欧洲人和亚洲人在此期间享受夏季，但仅在12 500年之后，相应的时间段将变为北半球的冬季。

所有这3种长期的因素——地球轨道的周期性收缩、旋转轴角度的变化和旋转轴方位的变化——一起构成了一种气候“太极”，从而影响地球接受太阳热量的多少和这些热量在地球表面如何按地理和季节分布。这些周期性的变化进行得很慢，要数万年的时间才能完成一个周期。在这种背景下，一个世纪内所出现的重要气候改变，例如20世纪由于人类活动而增强的温室效应，确实可以看作是非常快的。

即使更慢些的变化也会改变地球和它的气候，这些变化不是发生于数万年内，而是发生于数千万年内。在这一非常缓慢的改变的背景下，发生在仅仅数万年内的事物看起来是很快的！这一速度上的差异确切地说就如同一秒和千分之一秒之间的差异一样。当上述差异应用于曝光时，摄影师将喜欢这一差异。一个千分之一秒的曝光能够抓住活动中的运动员一个清晰图像，然而一秒的曝光将只能获得一个模糊的印象。刚才我们描述为太极的过程对于变化更加缓慢的背景而言就变成了闪电般的武术动作。

重新排列家具

什么是影响地球气候真正以蜗牛般速度发展的进程呢？板块构造和大陆漂移(continental drift)。作为地球内部大规模运动的反应，岩石构成的外壳碎片也在慢慢地移动(这些岩石构成了地球的固体表面)，这样地球表面的地形不断地被重排。这些巨大的镶嵌地砖被动地在内部流上方运动，其速度仅为每年一两英寸(几厘米)，大约是指甲生长的速度。如今美洲正以这样缓慢的速度远离欧洲和非

洲，大西洋正在逐年、逐世纪、逐千年地变得更宽了。如果我们看1000万年前的地图，当时处在这种慢速分离的早期，就会发现那时的大西洋比今天大约窄500公里。而且在2亿年前，大西洋根本就不存在，人们不沾一滴含盐的水就能从南美洲走到非洲。

地表的地形结构决定着洋流的模式。海洋占据了低的地方，大陆则占据了它上面的所有部分。大陆的位置决定了洋流能够到达和不能到达的地方。大陆地形最近的主要变化，至少就海洋洋流变化来说，发生在大约300万年前。当时巴拿马地峡上升，切断了大西洋和太平洋之间的联系。在赤道地区不再有一个东西向的环流，大西洋里洋流的流向主要变成南北模式：一个是发生在表层向北的暖流——墨西哥湾暖流；一个是海洋底部向南的寒流。它们一起形成了一个巨大的流体传送带，其表面的主要货物是热。

由于受北上穿越赤道到达北方高纬度区的大西洋的墨西哥湾暖流的影响，加拿大沿海地区、格陵兰岛南部、冰岛、英国和挪威的气候就远比北半球同纬度但远离大海的其他地方如加拿大中部或西伯利亚要暖和得多。甚至位于北极圈内很北处的俄罗斯港口摩尔曼斯克，由于受到围绕挪威极北点进入巴伦支海的墨西哥湾暖流的影响而终年不冻。如果海洋环流模式改变了，气候也一定会随之改变。

在巨大气候调节快速进行的时候，其他缓慢演化的地形重构也发生了。在南半球，大约300万年前围绕南极洲大陆开辟的一条连续的海洋路线，允许从西向东的洋流形成。这一洋流阻碍了温暖的水进入南部的高纬度区，有效地在气候上隔离了南极洲。随之出现了南极洲冰帽的累积，而且一直持续到现在。尽管地理学家对这些气候变化十分着迷，但这些变化对于评估20世纪气候变化的意义仅有边缘相关性。这些近来的变化必须根据地质年代上的最近全球气候系

统背景特征进行评估，在这个时间间隔内海洋洋流大体呈现代形式。这一时间间隔本质上就是自巴拿马地峡消失以后的300万年。

在某些方面，这过去的300万年可以被看作是一个“已打破的记录”。在那一时间间隔，洋底累积了层层的沉积物，其中包含微小的海底生物体的壳。这些壳讲述了约30次大规模的大陆冰河形成的周而复始的故事。由巴拿马包围而形成的洋流循环模式，以一种尚未被完全理解的方式，为周期性的冰河形成奠定了基础。但是在海洋底部的沉积层发现的冰河重复形成的记录是显而易见的。

海洋中的沉积物是如何讲述冰河形成的故事的？海洋中各种化学物质的含量取决于海洋中水的多少——水越多导致化学成分的浓度越低；水越少，浓度越大。生物体壳的化学组成反映了这一生物体存活时海水的化学组成。那些冻结成大冰原的水来自海洋，蒸发、流动，然后以雪的形式落到大陆上。当大陆上有更多的冰时，海洋中将有更少的水，反之亦然。最近的冰河期，在大约20 000年前达到最大程度，也只不过是这一反复进行的交响乐的最后章节。从那以后，冰开始退融，虽然并不总是以一种平滑和稳定的后退方式进行。

在理解地球气候的自然变化中，地质记录的确能为我们提供一些素材。但是这些经由过去的方法也让人们产生了很多的曲解。我们试图往回看得越久远，图像就变得越模糊。观察资料变得更少，在地球上的分布变得更不规律，解决短期时间间隔的能力变得更困难，确定某事是多长时间以前发生的就更不准确。在对过去2000年的气候变化的讨论中，人们认为科学文献中有一些根深蒂固的气候学概念需要重新评估。有关气候学的入门教科书写到中世纪变暖时期(MWP，约800～1200)和小冰河时代(LIA，约1300～1850)是由于全球气候的偏移。MWP是指斯堪的纳维亚人在冰岛和格陵兰岛定居

及其文化和贸易在欧洲扩展的这段时期；与此相反，LIA是指气候恶化的时期，在这个时期内斯堪的纳维亚人放弃了他们在格陵兰岛的定居点，冰岛周围冬季海冰区域极大扩展，欧洲的高山冰川不断增加，农业生产下降。[3]

就像我们也许很容易认为我们在某地区经历的事情是全球的特征一样，当人们考虑过去遗留下的迹象时，也容易陷入同样的错误逻辑中去。现在一些科学家怀疑，MWP和LIA是否发生在全球范围内。这一怀疑起因于上面提到的证据以欧洲为中心的特性。因为许多的历史记录是欧洲的，因此很有可能是欧洲人记录了这两个事件的影响，MWP和LIA只是主要影响北大西洋地区的区域性气候现象。如今我们对于北大西洋气候的敏感性以及它和北冰洋所发生的紧密联合了解得更多。除了通过遥远的北大西洋，北冰洋与世界上其他海洋并没有什么有效连接。在遥远的北大西洋，北冰洋的变化能够对与北大西洋接壤的陆地产生显著影响。但是在北大西洋地区有很好记载的那些影响，就构成了一个全球气候的变化吗？或者它们就是“把重要的地区气候事件曲解成全球变化”的很好例子吗？

这些问题回答起来不是很容易。即使MWP和LIA在北大西洋地区有自己的起源和重要影响，那里发生的变化也许能够以缓和的形式，通过大气流和海洋洋流模式传到全球的其他地区。对于地球某一部分的气候变化是如何与其他地方的变化联系起来的问题，还存在着相当大的不确定性。然而我们通过地区有限的观察资料做全球性解释时必须谨慎。一个全球性的描述通常需要全球性的观察。

小变化能够产生大影响吗？

人类除了很难察觉全球广大区域缓慢发生的变化之外，对全球气

候变化重要性的怀疑还有另一个来源：所经历或预测的变化的大小。当气候学家指出20世纪地球的年平均温度增长约1 ℃(差不多2 ℉)或者在21世纪还将增加几度时，询问“这样的变化是否重要”就很合理。毕竟，昼夜间的变化和季节性变化与之相比，变化要大得多。

回答以上问题的答案，部分在于生物通常能够适应相当大的短期变化，不会受到严重伤害这一事实。生理学机制已经进化到能让大大小小的动植物适应或者至少能忍受短期极端的状况，只要明天或下月或明年状况能够减轻一些。因此地球生物区的各个组成部分即使没有食物也能坚持一天，或者特别冷的情况下维持一周，或者降雨少的情况下支撑一年。但是我们都能够想象到一年没有食物，或者10年干旱的后果。生物体系所具有的保护性机理并不具有无限的弹性。即使干扰很小，因为小的影响会随着时间积累，它们也是能够被拉长到转换点的。我们都听过“压垮骆驼背的稻草”这个谚语，即使最后添加在骆驼上的负担是很小的，但这累积的后果是灾难性的。总有个临界点，也就是承受的限度，一旦超过了这个限度，不管达到这一限度的速度是如何慢或如何递增，最终将导致一个系统的崩溃。我们开始讨论全球性气候系统是否有可能到达一些转换点的问题，在后面一章我将重新回到这个话题上。

许多复杂系统对于小的变化非常敏感。大多数读者会很熟悉AM和FM电台波段。当上下班开车时，他们可以把收音机调到他们最喜爱的新闻报道、脱口秀或体育新闻上。我习惯将收音机调到89.1兆赫，它是服务我所在社区的几个公共电台中的一个。频率的标示意味着电台信号是由以每秒89.1兆赫传播的电磁波组成。然而，如果稍微调动一下到89.2兆赫，国家公共广播电台则被噪声取代。调到89.7兆赫可以听见古典音乐。百万波段中的一个小变化

却意义非凡，因为收音机被设计得要对细微变化很灵敏。

人体提供了一个很好的例子，这一复杂体系构造要在某个固定的温度下才能正常运转，即 37 ℃(98.6 ℉)。如果我们体内温度超过正常温度 1 度左右，这就说明我们病了，人体系统遇到了一些问题。1 度在理论上可能看不出有多大变化，但是对于一个调控良好的系统，如拥有对温度敏感的器官和构造精巧的温度自动调节器的人体，一个小的偏差都不是件小事情。

一个小的改变对于全球气候系统也不是件小事情。全球气候系统也是一个相互联系、相互交织的复杂体系，它吸收或反射太阳光，通过大气和海洋系统循环传输热，在系统内部进行化学物质的交换。系统的一个部分遭到破坏，系统的其余部分也会受到影响。在这个极其平衡的全球气候系统中，如果 1 度既是全球性的也是长期性的，则意义非凡。

关于进化的思考

本章所谈到的很多问题，在著名的生物学家威尔逊(Edward O. Wilson)* 最近出版的书中也都有所提及。在《生命的未来》**[4] 中他这样写道：

人脑的确进化到在情绪上对一小片土地、有限的亲属和未来

---

* 威尔逊(1929～)，美国著名生物学家，美国总统的长期顾问。在北美和欧洲拥有 27 个名誉博士学位，先后获得 75 个具有国际影响的奖项。他的著作《论人性》(1978 年)、《蚂蚁》(1990 年)曾先后荣获普利策奖。——译者

** 《生命的未来》主要是以一个生物学家的立场来论述生物多样性问题的，大量扎实的数据和例证，再加上亲身的研究和体会，他确凿地论证了生物多样性丧失的危险，并提出了解决危机的若干方案和办法。详见《生命的未来》，陈家宽、李博译，上海人民出版社，2003 年 1 月。——译者

的两三代负责。这种眼光既不太超前也不太离开家乡，在达尔文（Charles Darwin）的意义上是初步的……为什么我们眼光如此短浅呢？原因很简单：它是我们旧石器时代遗留下来的固定不变的部分。

如果威尔逊的观点“进化是根植在大脑中的空间和时间的乡土观念”，确实是正确的，那么我们也许怀疑，为什么自然没有选择把更多的先代积累的知识赋予给每代新人呢？思想家桑塔亚纳（George Santayana）* 忠告道，“凡是忘掉过去的人注定要重蹈覆辙。”[5] 为什么每一代的孩子们必须进学校学习基础算术？为什么我们一生下来不把乘法表根植在大脑呢？为什么几个世纪以来科学与工程系的学生还要重新学习牛顿（Isaac Newton）和莱布尼茨（Gottfried Leibniz）在 17 世纪就发明的微积分？如果孩子们天生就具有这些有力而不变的工具，而不是一次次地努力获得它们，这样不是更有效率吗？即使忘记了至少还可以记住一些。抛弃先辈们遗留下来的许多东西，又不断对下一代人灌输已获得的知识，这种进化策略的优势何在？

当然，大自然的确教给了我们基本的原理。我们的基因赋予了我们身体和大脑：身体提供了可动性和感知器官，通过它们我们可以获得信息；大脑则是控制中心和数据库。几乎每个人都由具备这些必需要素的生产线生产出来。接着学习便开始了。就像前面刚刚提到的，孩子们生来是具有好奇心的。他们使用自己的感知器官，去观察、去触摸、去倾听、去品尝。在他们的脑海里，他们分类、比

* 桑塔亚纳（1863～1952），美国哲学家。主要著作有《怀疑与动物信仰》、《存在王国》、《理智的人生》。——译者

较、评价这些经历。最后他们去学校里学习乘法表，有一些人甚至去学微积分。

没有把算术根植在我们脑子里的代价就是，我们每个人和每代人都必须重新学习加、减、乘、除。但是强迫每代人了解世界就是激发每代人的创造性的自然方式，这可行吗？也许自然已经选择不用太多的传统智慧去偏袒每一代人。世界总是在变化，人类不管喜欢与否，也必须进行改变。他们面对或适应未来变化的方式不会在过去的传统智慧中找到。这是威尔逊所解释的困境：过去的教训阻碍我们适应未来的困难。

自然没有给我们施加沉重的包袱，这些包袱会抑制创造性并妨碍我们顺利朝着未来发展，这对我们有益吗？我们脑子里熟悉的“十进制”算法可能会妨碍我们对现代数字计算机的“二进制”算法的概念化吗？无论我们怎么努力，我们都不能让世界静止。变化是正常的，也许承认和适应变化的能力在人类发展进程中比集中在传统智慧和保持现状的其他技能更为重要。

在本章中我们已经看到，诸如全球变暖这种现象是很难在个人经验的基础上进行评估的。我们可以很好地理解我们周围短时间内发生的事，但要了解全球平均温度的长期变化趋势却难得多。长期全球现象无论是在时间上还是在空间上都超越了我们个人感知的能力范围。我们只能通过世界各地几个世纪以来搜集到的观测资料和共享的信息来推导这一现象。接受和确信这些集体经验的精髓需要我们超越个人经验，迈入这个不熟悉和不确定的领域。谨慎地采取措施是自然的，对不熟悉世界的综合概括遇到一些怀疑也是自然的。

在下一章中，我们将着手讨论测量和数据集合的不确定性，这是

学科学的学生在迈进实验室时会面临的问题。有人曾经说过："如果你不能测量它，那么它就不是科学。"但用什么工具来测量？又有多少与之相联系的不确定性？在我们进行了一些测量之后，怎么来处理测量数据？不确定性花园的这部分是一个更悠久、发展更好的领域，它提供了一些与当代有关的颇为有趣的历史性观点。

注　释：

1. 长期趋势如何在不知不觉中慢慢走进我们生活，是 Hal Kane，*Triumph of the Mundane: The Unseen Trends that Shape our Lives and Environment* 一书的主题，Island Press，2001，200 pp。

2. Texas Transportation Institute，Texas A&M University，2001。亦可参见 the *New York Times*，8 and 9 May 2001。

3. Brian Fagan 的著作，*The Little Ice Age: How Climate Made History 1300 ~ 1850* Basic Books，New York，2000 (246 pp.)提供了小冰期日常生活中引人入胜的细节。

4. 2002 年由 Random House，Inc 的子公司 Alfred A. Knopf 出版。

5. George Santayana，*The Life of Reason* (vol. 1)，1905.

第五章　发烧还是着凉?

我们试图用码尺测量细菌。

——保罗斯(John A. Paulos)教授，[1]坦普尔大学

从基础水平上讲，当我们进行测量时，科学不确定性就开始了。我们用什么来进行测量? 测量工具如何能够很好地完成测量? 我们测量某一物体的尺寸、质量或温度能有多高的准确度? 如果我们多次重复一个测量，这些单次的测量结果彼此间有多大的一致性?

保罗斯教授对细菌测量发表的评论是在一种不同寻常的背景下进行的，对此我将在后边加以详谈。评论强调选择与手头工作相适应的测量仪器是非常重要的。一个人无需具备实验科学家的经历也能知道用一个码尺准确无误地测量出一个细菌的长度的可能性是非常小的。码尺最小的衡量标度通常为1/16英寸，虽然比2毫米小些，但比任何一个细菌都要大得多(实际约大出10 000倍)。因此一方面，人们除了知道细菌长度比1英寸的1/16小得多之外，可能也说不上

更多的什么了；另一方面，原则上讲，1码尺是可以测量首尾相连排列起来的500万个细菌的长度的。

绝大多数的物理或化学教科书都介绍了测量中的不确定性。这些讨论典型地集中在测量仪器的灵敏度（或精确度）和准确度以及测量同一物体或测量条件的可重复性方面。灵敏度与仪器对变化的反应以及测量仪器能够探测或觉察的数量差别的大小是相关的。准确度描述的是测量值与实际值的接近程度。温度计通过其液体柱高度的微小但可注意到的变化来表示1/10度的改变，而不是表示1/100度的改变，说明它具有1/10度的灵敏度。考虑以下的例子：沸水的温度按摄氏标准被定义为正好100℃。如果一个置于沸水中的温度计上显示的是99.7℃或100.3℃，那它就具有0.3℃的准确度或误差范围。另一种规定准确度的方法是采用百分比误差，在这个例子中就是0.3%。如果在一个特殊的实验中，化学家没必要知道实际温度测量的准确度在半度以内，那么不太精确的（也更便宜的）温度计就足够了。

如果你对同一物体测量两次或者更多次，直觉上你会认为每次的测量数据都是相同的。那么，不管是测量地毯的长度，还是用温度计测量地毯的温度，每次当我们对卷尺或温度计读数，并且当读数落在测量仪器的最小刻度之间时，我们都要把这个数字估读出来。由于观察角度和观察者的经验不同，因此每次对测量进行的估计可能都有所不同。同样地，如果我们计算某一血样中白细胞的数量，原则上无论是哪个实验室的技术员进行计数，人们会认为每次的计数是相同的。但事实上，每个实验室技术员的结果都可能互不一样，甚至同一名技术员每次结果可能也不同。在任何重复测量的数据集合中，都有一个值的范围。如果有的话，哪一个是正确的呢？我们经常依赖统计数据来确定平均值，并对从多次测量中计算出的平均值是

真实值的概率进行估计。如果我们进行多次测量，这种概率一般较高；如果仅测量几次，此概率会低一些。但是，这个概率绝不会达到完全符合真实结果的地步，依然存在着一些不确定性。

不同的人对同一事物进行测量时，重复测量中当然同样存在着上述问题。但他们使用的是不同的测量工具或不同的技术。例如：使用不同的实验方法确定光速的大小，光速就会有与实验方法相联系的一些不确定性。但不确定性会随着更敏感设备的设计和利用而变小。

可重复性是科学的一个基本原则。如果一位科学家宣布一项令人震惊的科学发现，其他的人将会试图通过重复实验或测量来证实这种现象。1989 年，当一项关于原子在低温下聚变[2]（一般称之为冷聚变）的实验公诸于世时，立即引起了人们的兴趣。如果冷聚变被证明实际是可能的，就有希望产生无限廉价的能量。世界各地的研究人员冲进实验室，亲自尝试这个实验。在这种情况下，可重复性变得难以理解，冷聚变的前景就像用过的罗马焰火筒（roman candle）落到地上一样。最近的关于冷聚变的声称以及第二章中提到的气泡内爆声称，相信世界各地的实验室会很快对它们进行实验。

由于要测量的量是变化的，因此并不是所有的测量都能重复。室外的温度计将会显示温度一直都在变化。购买旧房子的人发现墙与地板、天花板之间的夹角，随着时间的推移，由于居住和下垂会变得不再是 90 度。逐天看来，几何形状也许呈现静止状态，但很多年下来，变化就累积起来了。当地球表面的地壳由于板块构造力而向上提升以及因侵蚀而减小时，珠穆朗玛峰的高度也在改变。随着时间流逝，我们可以获得一系列测量数据，但由于时间在进展，我们永远不会有机会来重复许多早期的测量。

不准确的另一方面，因此具有的不确定性，在任何测量中都与测量仪器的刻度有关。我们怎么能知道卷尺或码尺或温度计上的标记是被均匀隔开的呢？我们能肯定这些测量工具上的1英尺是实际的1英尺距离吗？我们能确信体温计读数为98.6 ℉时，实际的温度真是这样吗？或者也许是98.5 ℉，或者甚至更糟，也许是102 ℉呢？温度计的制造商们是如何确定多高的液体柱实际代表98 ℉、99 ℉和100 ℉呢？答案就是：他们利用另一支“更好的”温度计来判断一个物体或液体的温度实际上到达98 ℉时，然后就在被校准的温度计里的液体水平面的外面玻璃上做个适当的标记。

但是这“更好的”温度计是如何有了标记的呢？讲到这，有人会马上认识到需要一个标准：我们假定的就是确切的1米长或确定的某一特定温度的一个状态。在美国，国家标准和技术协会对此有最终发言权。任何测量仪器的性能都可以和公认的标准值进行比较。那些测量数据与标准很相近的仪器要比那些测量数据与标准出入大的仪器好些（好的通常成本更高）。但是，除了读仪器时的不确定性以外，无论用哪种仪器，对于仪器测的是否是真实值也有一定程度的不确定性。

标准量度

美国第六任总统亚当斯（John Quincy Adams， 1825～1829）曾说过：

> 重量和尺寸也许被列入人类社会每一个人生活的必需品中。它们成为每个家庭经济安排和日常生活的一部分。

对于距离、时间、温度、质量和许多其他物理量的测量，已经建立了测量的标准。随着新技术方法的开发，这些标准也随时间而不断发展。“米”作为测量公制中长度的基本单位，其标准的发展历史，为人们了解物理量是如何定义以及标准是如何建立的，提供了极具吸引力的深刻见解。

标准“米”的故事开始于18世纪早期。在那时，科学界仍然惊讶于半个世纪前牛顿在力学上的发现。特别是他根据万有引力作用下简单的运动定律成功地描述了行星围绕太阳运动，这为他赢得了巨大的赞誉。但是他的力学定律并不仅仅局限于解释行星轨道。它们也与许多其他的自然现象有关，比如地球的形状等。在这方面，万有引力所起的作用是让地球上所有的质量都尽可能紧地吸引到万有引力的中心，从而压进尽可能小的体积里。其结果是地球和其他行星将被这一过程理想地定形为球体。

然而，如果这个球体也像所有行星那样围绕一个轴旋转，那么另一种力——旋转的离心力，开始发挥作用，行星结构就必须适应这个附加的力。牛顿力学预言行星的旋转会造成它在赤道向外膨胀一点和在两极点稍微向内收缩一点。在万有引力和离心力两者的平衡下，行星就形成了一个椭圆，其长轴在赤道平面，其短轴通过极点。

科学理论总是要经历测试和检验，以便检查其是否不完善或是否存在缺陷。牛顿力学也不例外。这样，对旋转地球形状的预言所进行的直接测量也是牛顿力学的另一个证实或证明。由此，在1734年，也就是牛顿逝世后不到10年，法国科学院构想出一个实验，此实验为牛顿建立的革命性物理学提供了一个极其重要的检验。法国派出地质学家和调查员去北斯堪的纳维亚和秘鲁测量地球曲率，而这些地方分别接近于极点和赤道。历经3年的野外调查证

实，地球的确在低纬度膨胀，在极地附近扁平。牛顿物理学通过了另一个重要检验。

那么，上述故事和建立测量标准，特别是与定义长度标准（我们称为米）有什么关系呢？作为这次地质考察的自然结果，法国科学院做出一个决定，利用地球本身而不是地球的某个部分，作为一个测量标准。他们决定沿着子午线把从赤道到极点的距离分成1千万等份，把每一单元长度称为1米。如此一来，米作为一种长度测量单位，最初的定义为沿着一个子午线从赤道到极点的距离的1/10 000 000。当然，他们实际上没有从赤道到极点摊开一个长卷尺测量距离。他们知道这个距离与90°纬度相符合。如果他们对高纬度和低纬度处的1°纬度的长度进行仔细的测量，那么通过简单的测量就能推出全部的长度。采用这种新标准，科学就不再依赖于某个国王脚的长度，或者其他相似的"非科学的"尺寸。科学家们自此以后有了一个量度标准，这个标准直接与永久的地球有关，而不是维系在会死亡的人类的尺寸上。最后这个地质测量规定被转化为一支铂铱合金棒上的标记，它保持在一个标准和不变的压力与温度下，以避免因环境改变而造成的尺寸轻微改变。这一金属棒存放在巴黎，被国际上一致认可为1米长的最终的物理参考。

尽管如此，多年以来测量学的改进和对地球尺寸更好的测量导致出现了与1734年法国确定的标准的偏离。今天，大地测量工作者们认为沿着子午线从赤道到极点的距离是10 002 286米，而不是10 000 000米。1734年最初的误差只有大约万分之二，或者每米大约0.2毫米误差。这对那个年代来说已经是相当准确的测量了，其误差在1米中只有1根头发的宽度。很明智的是，我们没有根据对地球进行的这一更新和更精确的测量数据重新定义米。另外，随着

时间的推移，地球的自转逐渐变慢，赤道的膨胀和极地的扁平将逐渐地、不断地调整，这使得地球的尺寸也一直在变动着，结果对国际测量标准来讲也不利。

最近对米的定义实际上已经完全脱离了地球表面形状，而倾向于赞成更加稳定的、从不变化的事物。1960 年，根据从受激原子氪 86 所发出的电磁辐射中的光谱线中某一特定谱线的波长的特定倍数(确切值是 1 650 763.73 倍)，米被重新定义了。米不再直接地与地球的几何形状有关。[3]然而，1960 年对米的重新定义的使用期是短暂的。1983 年，米又一次被重新定义。这一次是光在真空中 1/299 792 458 秒内所传播的距离。这个定义有效地说明了光在真空中的速度是每秒 299 792 458 米，同时也含蓄地承认了没有什么能比光速更能准确地描述米的标准了。[4]

类似的情况也会出现在测量系统其他基本单位的确立中。公制中的质量，最初也是法国人定义的，被称为克，其值相当于在某一特定温度下 1 立方厘米水的质量。然而，这一标准在物理上是一个非常小的量，大约有一小管水那么多，是很难精确测量的。今天这一量的标准是千克(1000 克)，由在巴黎国际度量衡局的金库中保存的一个圆柱形的铂铱合金(美国的国家标准和技术协会中也有一个复制品)代表。这个 1 千克圆柱体是保留的仅有的用实物人工制品代表质量的标准。时间的标准为秒，最初的定义为地球绕轴旋转 1 次时间的 1/86 400，旋转 1 次是 1 天，1 天是 24 小时，每小时包括 60 分钟，每分钟包括 60 秒。然而，现代对地球旋转更详细的测量表明地球的旋转不是固定不变的，因此作为时间标准也是不可靠的。今天，官方把秒定义为从铯 133 原子放射特定波长的电磁波周期的 9 192 631 770 倍。就像米的定义一样，秒的定义也与地球特征分开了，被更加稳

定的特殊原子行为所代替。

在大海中迷路

需要也许是发明之母，但不确定性的确是她的孪生姐妹。不确定性一直是创造性的一个刺激，就对时间进行测量的天文钟而言更是如此。某人在不熟悉的环境里，最常见的不确定表述之一就是：“我在哪里？”当水手们在大海中看不见为他们指明方位的路标时，他们对上述不确定性尤其深有体会。航海的整个艺术和科学起源于人们渴望知道自己在地球表面的位置。没有路标，海员用天上的恒星和行星这些“天空标志”来判断方位。但是由于地球每天都围绕轴旋转，天上的向导也并不是在天上原位不动，似乎也随着其下边地球的旋转而运动。但是有个例外，那就是北极星，北极星位于地球旋转轴延长线的北极点的正上方。在夜空中它看起来是静止的，观察者只要测量北极星与地平线之间的角度就可以知道自己目前所处的位置。如果你在北极点，北极星就在你头顶，或者是与地平线成 90 度角——这就是北极的纬度，北纬 90 度。北半球的水手很早以前就知道用北极星简便地确定纬度。

然而，经度也有另外一个故事。经度衡量离任意出发点的东西距离。今天国际上一致认为，经过英国格林威治的那条子午线为零经度，称为本初子午线(prime meridian)。但历史上也有其他有关本初子午线的选择，多年来法国把经过巴黎的子午线作为本初子午线。当一名海员测量地平线与天上某颗缓慢运动的星星(由于下方的地球在转动)间的角度时，判定某个位置的经度问题本质上相当于对时间进行仔细测量。太阳就是一个很好的例子——它升起，慢慢成曲线向上到达一个高点，接着开始下降一直到日落。如果知道日升、最高

点、日落的时间(或者是其中任何位置的时间)，而且知道是一年中的哪一天，就可以确定观察者离本初子午线的东西距离。其中的要诀就是只要在观察时有一个能够很好读取时间的好时钟就可以了。

在早期的探索中，大多数的时钟是机械装置，带有齿轮和发条。它们并不完善，速度易变，或快或慢。机械钟的偏差取决于当地温度、大气压和相对湿度、润滑油的自然特性、发条的松紧，以及它是水平还是垂直安放的等等。对于大海中的船只来讲，在海上航行数月而没有路标，准确判定船在地图上的正确方位或将新发现的陆地或暗礁标定在航海图上的能力，与准确地保持时间的能力是相关的。因此就需要设计和建造一个稳定和精确的航海天文钟来减小处于茫茫大海上哪一位置的不确定性。英国海军部对开发这样一种仪器有很浓厚的兴趣，并提供极其丰厚的奖金奖励发明满足海军部要求的航海天文钟的发明者。18 世纪的钟表制造商之间随之展开的竞争，以及以后海军部一直拖延不给获胜者奖励的故事，成为索贝尔(Dava Sobel)* 笔下一本引人入胜的书的主题。[5]

"鹰已着陆"

当宇航员阿姆斯特朗(Neil Armstrong)** 以清晰无误的话语"鹰已着陆"*** 宣布阿波罗 11 号宇宙飞船到达月球表面时，也许没有什么比收到这样的无线电广播更令人兴奋的了。然而，在早期的无线

---

* 索贝尔，美国作家。著有国际畅销书《经度》。2002 年 3 月，上海译文出版社出版的由谢延光译的名为《伽利略的女儿——科学、信仰和爱的历史回忆》的传记也出自她的笔下。——译者

** 阿姆斯特朗(1930～)，美国宇航员。1969 年，曾与奥尔德林(Buzz Aldrin)和科林斯(Mike Collins)驾驶阿波罗 11 号飞船完成了人类首次征服月球的壮举。——译者

*** 指阿波罗 11 号在月球成功着陆。1969 年 7 月 20 日，当登月舱在月球表面安全着陆时，宇航员阿姆斯特朗立即向地球报告："休斯顿，我们在月球静海报告，鹰已着陆！"——译者

电通信中，无线电广播发射机和接收器按现在的标准来看都是粗制滥造的东西。任何发消息给远距离接收器的人喜欢听到的话是“响亮和清楚的”，而不是“微弱和混乱的”。为了能让接收到的信号声音响亮和清楚，它就不能被其他的声音，比如静电噪声、噼啪声、嘶嘶声或嗡嗡声遮住。在传播时，信号也不能变形得听不懂。当其他声源干扰接收所需要的信号时，我们说我们处在“嘈杂的”环境下，很难听到“信号”。同样的描述可用于婚宴中，在那里充斥着舞曲声和其他人的闲聊声，你的声音在嘈杂声中很难听清楚。信号迷失在噪声中。相反的，警车警笛声或者一辆救护车的高音喇叭声则是有目的的：它们被设计成很容易就能从通常的交通噪声中辨别出来。

实验室的科学实验报告和测量仪器就好像一个无线电接收器，原因在于科学家只想测量关系重大利益的量，并把来自其他方面的不相关的数据最小化。由于避免所有的噪声是不可能的，当信号处于足够小的水平时，信号声将消失在噪声中。因此在设计科学仪器时，能够估计信号和噪声两者的大小是非常重要的，以便仪器可以满足敏感度和精确性方面的要求。

现在从广义上来讲，任何干扰信号的东西都可被称作噪声，它的来源是很多的。在树木气候学(通过研究树木的年轮，揭示树木存活期间当地的气候状况)领域，科学家分析某个地区中的许多树木，而不是单一的一棵树。他们这样做是因为任何单棵树也许只是显示了树所处区域特有的重要地区性噪声。一棵生长在流域盆地偏远地方的树也许比另一棵生长在水分充足的山谷中的树更容易受到降水量的影响。森林中的树木要为争取接受更多的阳光进行竞争，然而分布稀疏的树木可能就不会有这样的竞争。一场地震也许会阻碍树根汲取水分，但仅距几英里之遥，地震就不会产生这样的影响。简而言

之，没有单独的一棵树木能够可信地讲述一个可靠的故事。但是，一个地区许多树木组成的整体，其中每一棵树都受到所处位置特有的影响和地区气候变化的影响，能够让科学家把地区信号从所处位置特有的噪声中分离出来。

清点选票

保罗斯教授对用码尺试图测量细菌的评论不是在实验室中测量细菌的背景下，而是在候选人实力相近的选举中确定其结果的背景下作出的。*2000 年 11 月，美国人对于谁当选为美国总统，陷入了一场旷日持久的不确定性当中。

在大多数的民主选举中，确定谁获胜在概念上是简单的：清点选票，选票最多的候选人就是获胜者。佛罗里达州在决定谁将成为总统方面被证明是关键的一个州。两名主要候选人布什和戈尔(Al Gore)差不多均分了初始的近 600 万张选票。两名候选人的选票总数只差不到 1000 张，这意味着每 6000 张选票中，每名候选人的选票数仅差一票。自然，由于此事非同小可，因此明显落后的一方就要求重新计数来确保统计选票中没有发生错误。

任何曾试图统计大量事物的人，比方说清点一个大罐子里的硬币数目，都知道要得到一个完全精确的数字并不像看起来的那样简单。举一个例子，让我们想象一个装满 10 000 个绿色弹球和蓝色弹球的桶，我们想要知道桶中每种颜色弹球的数目(这是对两名候选人竞选投票的模拟计算)。如果我们把弹球倒在地板上，然后根据颜色将其

---

* 参见《我思故我笑——哲学的幽默一面》，约翰·艾伦·保罗斯著，徐向东译，上海科技教育出版社，2002 年。——译者

分开并进行统计，不确定的结果常常会出现。也许是计数时不注意、分心、错认或者是计算错误，不同的人每次在数蓝色和绿色弹球时得到的最终数目可能会有所差异。

如果仅仅是确定绿色弹球是否比蓝色弹球多，那么通常计数结果之间的差异就不重要了。如果在10 000个弹球中，实际上有6000个蓝色弹球和4000个绿色弹球，那么，如果一个人计数的是5984个蓝色弹球和4016个绿色弹球，而另一个人数的是6007个蓝色弹球和3993个绿色弹球，这并不重要。即使第三个人的数目是5991个蓝色弹球和4007个绿色弹球，总共加起来只有9998个弹球(也许有两个滚到附近的橱柜下而没有被看到)，这也不重要。就算有人得到的数目超过了10 000个也没关系(也许是一些弹球碎成了两半，而在计算的时候把每一半都包括在内了)。也许因为染料的变化问题或因为计数的人是色盲，甚至会存在着难以分辨弹球是蓝色或绿色的困难。

所有这些不确定因素都没有让人对“蓝色弹球明显多于绿色弹球”这样一个结果产生怀疑；蓝色候选人赢得了选举，绿色候选人则失败了。这是因为不确定性的范围比候选者选票总数间的差异小得多。让我们加上或减去10张选票并把计数看作是精确的(这意味着最终的计数会因增加或减少10张选票出现错误)，由于蓝色候选人和绿色候选人大约有2000张选票的差异，20张选票的不确定性范围只占差异的1%。信号比混乱的噪声要清晰得多，结果中的不确定性显然并不重要。信号声音大且清晰：蓝色赢了。然而，如果桶里的弹球实际上在颜色上是平分的，即分别有5000个蓝色弹球和5000个绿色弹球，那么在每次计数中，20个弹球的不确定性就很容易产生不同的获胜者。测量工具，即计数弹球的总过程，没有判定蓝色弹球和绿色弹球是相等的所需要的准确性。

在2000年美国总统大选中，差不多600万张选票中，两位候选人各自赢得的选票在佛罗里达州只有几百张的差距。关于一张选票是否包括在最后计数中的不确定性的主要来源集中在选举程序上。选举程序是，选民从一张所有候选者列表的投票卡上，通过在其中一位候选人的选票上打一个小孔，即孔屑来进行投票。然而有些孔屑还是粘在一起，并没有完全分离。吊在一个连接点上的孔屑能被统计进去吗？吊在两个或三个连接点上的孔屑又如何呢？或者根本就没有发生粘连，但孔屑呈锯齿状，也许表明投票者已经打算投票选那个候选人，而只是在使用所提供的工具时把孔屑从投票中分离出来没有成功呢？

程序的其他方面也让人对一些缺席选举人实际投票的正确性产生怀疑。缺席选票申请正确填写了吗？缺席选举人的选票是在截止日期前邮寄并接收到的吗？一些缺席选举人的选票被放错地方但后来又发现了。一些投票者对设计不同的其他选票感到迷惑从而导致没注意就把选票投给不同的候选人，而不是选举者想要选举的那个人。而且少数一些选票可能是由一些没有资格的投票者、一些未登记的人或一些非公民投的，也许还有一些投票者甚至在不同的选区投了两次票。也有人宣称一些合法选民无法参加投票，因为他们的名字被错误地从登记投票者的名单中删掉了。

在佛罗里达州出现的投票和计票过程中的缺陷并不是唯一的。它们会多多少少、不同程度地出现在各地的选举中。选举过程中出现这些问题是不可避免的，但值得欣慰的是这些缺陷通常并不很重要。我们在确定选举获胜者中所用的测量工具已经够精确了。但是在实力十分接近的选举中，由于候选人选票的差异和测量的不准确度大小相同，那么内在的不精确性就会凸显出来。这种差异正好在选

举过程和计票过程中的不确定性范围之内，因此选举结果也就陷入选举过程的不确定性之中。 佛罗里达州的投票和计票过程，以及其他大多数选区，面对选票总数中这么小的差距，缺乏足够的准确度以确定谁是获胜者。 在佛罗里达州投票和计票过程中，选举双方势均力敌。 一些报纸后来进行的重新计票支持了这个结论。 这个媒体协会认为结果取决于如何清点选票。[6]在更严格的解释标准下，对竞选者得票数的检查会产生一个获胜者；在不太严格的解释标准下，会产生另一个获胜者。

当然，正常的投票和计票程序不能确定选举的获胜者，并不意味着没人就职，而只是说明就职的人是通过其他方式选出来的。 美国人确信在宪法规定的日子他们肯定会有一个新总统宣誓就职。 因此当选举了无结果时后备程序就会发挥作用，这包括立法和司法的干预。 立法的解决方案仅是承认选举无结果，并取代立法机构的选举，无论此立法机构是州立法机关还是美国国会。 法院的干预是针对“灰色”选票，例如那些没有完整分离的孔屑所代表的选票或邮戳不清晰的选票，是批准还是否决得作出判决。

在2000年竞争激烈的总统选举中，美国最高法院仅仅通过以下方式有效地选举了新总统：在佛罗里达州的重新计票中，当布什领先几百票时，就命令该重新计票必须停止。 不确定性在选择总统的过程中并不是无限存在，宪法规定的程序明确规定面对不确定性时必须做出一个决定。 根本没有无限的“研究这个问题”的选择余地，尽管许多人可能会提出争议，他们认为还有充足的时间来对最初由机器数出来的选票进行人工检查。

美国2000年的总统选举提出了一个问题，就是我们是否能够设计出更好的计量工具，这样就能够得到一个更完整和更精确的投票

数。实力十分接近的选举的不确定性迫使投票者和他们的州长以及立法者去思考怎样才能改进投票和计票的程序；不确定性已经激励了创造性和灵活性。如果所有的选举产生的获胜者拥有 60%的选票，目前的标准是足够了。但在 2000 年的总统选举中是不够的，因为当时每名候选人获得了大约 50%的选票，相差不过 0.0001%。这次国家因为测量仪器的不完善所经历的事情促使人们就如何改进美国的选举程序提出了许多建议。

在测量之外

绝大多数科学和统计学教科书中关于不确定性的讨论都结束于处理测量、标度和重复的不确定性。尽管这些不确定性的概念很重要，但事实上，在整个科学领域它们并不是不确定性的最大根源。不确定性的更大问题来源于远在测量工具的精确性和准确性之外的其他方面。其中的一个问题集中在我们怎么样才能把一次的观察集中到一个更大的图景中。法国艺术家修拉(Georges Seurat, 1859～1891)* 创立的绘画的点彩派** 风格，非常接近于色彩涂抹的集合。从远处看，这些点就变成了有代表性的图画和艺术杰作。这个问题简化为："对于森林来说，在单棵树木测量基础上建立森林的数据集，还有比这更重要的吗？"

1998 年初，一家报纸的标题宣称"1997——历史记载中最暖的一

* 修拉，法国画家，后期印象画派代表人物。他把文艺复兴传统的古典结构和印象主义的色彩试验结合起来。把最新的绘画空间概念、传统的幻象透视空间，以及在色彩和光线的知觉方面的最新科学发现结合起来。对 20 世纪几何抽象艺术有很大的影响。——译者

** 点彩派是法国新印象派的一个分支，在这一分支中视觉混频或碎色的原则被推向极端——即以细小的点或小而分散的笔触来上色。只有隔开一段距离，当观察者通过视觉作用将那些色彩斑点调和起来而产生可见物体轮廓时，点彩派画中的人或物的形状方可见到。点彩派的发明者和主要代表人物是修拉，另一领导人是西涅克。——译者

年”。[7]读过这篇文章，人们就会了解到有一项对数以百万计的温度测量数据的分析，这些数据由世界各地的气象站以及海洋上的浮标和船只所测得。这项分析表明，1997年的全球平均温度超过自进行系统化温度测量和搜集以来的任何其他年份。但是全球年平均温度究竟意味着什么呢？平均温度是如何被确定的？为什么在这样的数字上也存在不确定性呢？

让我们先从测量某一给定时间和给定地点的温度开始吧。你可以很容易想象一个人早上醒来，透过窗户扫视室外的温度计，希望得到今天该穿什么衣服的线索。假如这个人每天都记下温度，到年底时把所有的读数平均一下你就可以判定当年早晨的平均温度。但是，偶尔我们这位认真的观察者忘记了读数或者在读数以前没有戴眼镜，或者是计数有误，或者去度假了，或者不小心打破了温度计，所有这些过错和疏忽都导致了数据差距或数据错误。此外，假如要换温度计，或许我们的观察者会决定买一个不同类型的温度计，它有着不同的灵敏度和精确度。因而对于同样的温度它的读数和原来温度计的读数会稍有不同。而且我们应该如何比较室内阴影处的温度计显示的温度和隔壁向阳处的温度计显示的温度呢？所有的这些问题都可以提出和调和，但每个问题让确定那一地点的年平均温度都增加了一点不确定性。

确定某一地点的年平均温度的另外一个考虑，是和以下两个问题相关的：如何选择温度测量的时间；如何平均不同时间测得的温度。一个人每天早晨当他拿起报纸时勤奋地观察并记录温度读数，这样经过整整一年，把这些观测值进行平均，就可以得到年平均值。但它可能是上午6点的年平均值，而不是代表24小时的年平均值。我们的观察者或许还会记录下下午6点回家时的温度，并根据早上和晚上

的温度估计全天的平均温度。除了早上和晚上观察外，如果他还在中午和午夜记录温度，那就做得更好了。你看到了关键所在：要获得某一地点的较正确的日平均温度，必须通过代表在白天和夜晚间摆动的所有阶段的温度的观察来确定。一旦我们很好地估计了某地日平均温度，我们原则上能够很容易地推出按月算和按年算的温度平均值。

当然，在现实世界里，我们并不靠业余观察者每天读取室外温度计的读数。这项工作是由工作在官方气象站的专业气象学家承担的。在这样的气象站里，测量温度是主要的工作；在一个标准的房间里放着一个标准的温度计，温度至少每小时记录一次，有时是连续记录。在理解过去几十年甚至几个世纪气候变化的原因方面，逐小时的详细信息是非常珍贵的，因为它对于确定“是平常的白天暖和的夜晚或是暖和的白天平常的夜晚引起变暖，还是变暖影响了白天和晚上的温度”这个问题的答案是非常重要的。每一种模式的产生都是对不同的气候刺激的反应。正如我们假想的业余观察者一样，气象站偶尔也会出现温度计受损、动力故障而短时间停电、因为资金不足和战争而关闭，以及从城市迁址到郊区等各种情况，所有这些都会造成数据差距和数据错误，从而导致测量的不确定性。

确定全球年平均温度的不确定性的另一个来源与温度的测量地有关。并不是在地球表面的每一点都进行温度测量，而仅仅是对全球不规则分布的几千个点进行收集。例如，欧洲和美国的气象站要比西伯利亚、亚马孙河盆地、撒哈拉沙漠、南极洲的气象站要多得多，而后者中任何一个地域都和欧洲或美国差不多大小或更大一些。为了便于讨论，让我们设想欧洲有 400 个气象站，而西伯利亚、亚马孙、南极洲、撒哈拉加起来总共有 100 个。如果我们把所有这些气

象站的读数简单地直接平均一下，我们得到的结果就会在很大程度上向欧洲大量气象站的读数倾斜，这一平均值由于在欧洲具有许多的测量数据当然就会产生地区性偏差。很明显，人们必须很谨慎地不能让地球上某区域的大量观察数据在对整个地球平均值的计算上占支配地位。相对于那些来自易于观察的区域的测量值，人们必须让来自边远地区的较少测量值在平均时占有相当的比重。一个更好的办法就是把地球表面等分成许多区域，确定每一区域的平均温度，然后从每一份区域的平均温度确定全球的平均温度。但是很显然如果某地区的数据只是从几个气象站获得的，而另外的区域则是从成百个气象站获得的，那么由少量观察数据得出的地区平均值要比由更多观察数据得出的地区平均值有更大的不确定性。决定于区域平均值的全球值将反映与区域平均值相关的变化的不确定性。

与观察地点的地理分布相联系的问题，使得在 1860 年以前不可能直接根据温度计的读数来确定全球平均温度。其中的原因很简单：南半球没有建立足够多的气象站来测量在全球平均温度中代表那一巨大区域的温度。自从大约 1860 年以来，有足够的气象站分布到全球的很多地方，从而使得确定代表每一区域的平均温度成为可能。而且通过集合，也可以得出全球的平均温度。每一地区平均温度都带有自身的一些不确定性，从而给确定全球平均温度增加一些不确定性。

统计人口

大规模数据收集和综合的其他例子也可以帮助我们想象大量观察数据集合和分析中固有的复杂性和不确定性。例如，每 10 年宪法授权的全国人口普查都会提出“美国有多少居住人口”的问题。像温

度一样，人口数量也是易变的对象，它不是静止不变的，而且一直在变化。

进行人口普查，就像测量温度一样，也有许多不确定性来源。进行人口普查的一个中心战略是确定国家的每一居住地居民的数目。因此，通过邮件，把人口调查问卷送到每一位居民住处，询问基本的人口统计问题。但是，一些接受者会直接把表格扔进字纸篓，有些人也许会忘记填写表格。一些答卷者因为担心他们违反了一些居住标准的规章或是他们知道一些居民是非法移民，而不会如实地回答问题。一些人住在非住宅区。还有一些人，例如无家可归者，根本就没有地址。

人口普查中的少计并不是人口普查中出现错误或不确定性的唯一来源。人口多计的情况也是存在的。父母离异的孩子有时会在父母双方的表格中都被登记。寄宿学校的学生或者大学生有时会在家和学校都被统计了。也有同一家庭多次填表的可能。在2000年的人口普查中，表格被邮往从大量不同的表格中搜集到的地址。为了尽可能的全面，人口调查局把最初的邮件发送清单与来自美国邮政部门和地方政府的信息进行了合并。尽管进行了努力来排除重复邮寄，但肯定还是有疏忽。表格在政府办公室和便利超市随手可得。通过电话或因特网也可以进行人口普查，因此也很有可能出现重复统计。

通过电话或人口调查员走访乡邻的跟踪调查，会略微改进数据的收集，但在最后的分析中，人口普查结果依然还是不完全和不准确的，因此也是不确定的。就像直接清点选票并不能总是出现清晰的选举结果一样，直接点查全国每个人的人口普查战略也是有缺点的。

2000年美国人口普查的最终结果是人口大约为281 000 000，这是除去由于重复或多次统计而造成的多出的约3 600 000人口之后而

得到的数字。人口调查局估计实际低估的人数可能在 2 700 000 到 4 000 000 之间，或者占清点的人口的大约 1%～1.4%之间。而且就像在选举投票统计中一样，人口普查的缺陷并不是均匀分布在所有人口区域。最后人口普查估计中的不确定性也是不均匀地来自不同区域和不同人口。对西班牙后裔的低估率在 2.0%～3.5%之间，美裔非洲人大约为 1.6%～2.7%，印第安人是 2.8%～6.7%，18 岁以下儿童大约是 1.2%～1.8%。[8]下一章我们将分析为了提高准确性，如何使用统计方法改善人口普查结果。但是潜在的改进也涉及政治的卷入，而这可能会延缓更精确的人口普查。

读出言外之意

让我们转到从事聚集观察数据引起的实际问题上去。例如，地震发生后，考虑一下确定有多广的区域经受超过某一特定水平的地面加速度的问题。以下现象很有代表性：在震中附近，地面震动最大；离震中越远，震动就越小。地震周围经受大的加速度的区域有多大呢？对此问题的回答非常重要，因为对于在地震危险区确立学校、医院和核电站的建设标准，它扮演着非常重要的角色。与这项任务相关的测量仪器是一种被称为加速度计的仪器。这些仪器在地震可能发生的地区广泛使用，而且在诸如洛杉矶这样坐落于高地震危险区的一些城市中数目可达数百个。在远离这些地区的其他区域也使用这种仪器，但数目要少多了。洛杉矶的加速度计可能比整个北达科他州的还要多。

这些仪器并不全是来自同一个制造商，甚至设计也不相同。但是通过校准，地震学家认为它们完全可以确定地震附近或者几千公里外的地区的加速度（在可接受的不确定性范围之内）。因此地震过后，地

震学家把所有的读数汇总到一起，把它们标记到地图上仪器所在的位置。接着困难就来了：从离地震不同距离和方向上获得的零散读数的集合，科学家是如何估计经受超过某一确定值的加速度的区域呢?

让我们把0.1g的加速度(重力加速度的十分之一)作为影响值。在一个加速度计上我们看到的是0.12g，明显超过我们刚选定的级别。在另一个距离地震更远处的一个站点记录的读数是0.03g，这又明显低于选定的标准。事实上，在所有的仪器中有些标记高于0.1g，而有些低于0.1g，也许没有任何仪器读数正好在0.1g上。地震学家面对的任务是通过这些不规则分布的点画一条线，从而把超过0.1g和低于0.1g的值分开。实际上，这条线是地震学家对仪器记录的加速度值为0.1g的地点的一个最好估计。一旦画出了这条线，就很容易确定这条线所包围的区域，但是在那些拥有仪器读数的地点之间恰当地画出这条线，则需要对很多判断进行检验。

估计一个给定的值处于两个已知值间的哪一位置的过程被称为**内插法**，类似于当玻璃管中的液体顶端处于玻璃上的刻度之间时读取温度计的读数。在温度计的例子中，我们通常认为超过较低标记的温度增量直接与包括液体柱顶端的两个刻度间的距离成比例。对于地震，也许有人认为加速度减小的线正好是离震中距离增加的那些圈，就像石块投掷到池塘，水面泛起的波纹向外扩散一样。但是地震变化的模式决不像几何学那样简单。地面加速度对地球表面土壤或岩石的类型以及它是否渗透了水非常敏感。震级也受地震与记录仪器之间亚表层的地质结构的影响。几十公里内可能造成两个仪器的读数不同，而且所有这些因素在两个地区之间也有很大的不同。

确定经受某一特定地面加速度的地区所具有的不确定性会带来进一步的后果。为了便于讨论，假定引起上述加速度的震级是6级，

早期的研究确定经受同样加速度的更小地区是5级地震和4级地震引起的。每种情况都因为上面提到的同样原因存在着不确定性。所有这些调查结果表明了当地震级别增加时，更广的区域感受到了这样的加速度。运用这些结果，我们现在就想估计有多大的区域受到如震级是7级或8级这样真正的大地震同样的震动。很明显，在确定受3个更小地震影响的地区中出现的不确定性，会对“由于受到更大级别的地震的影响而发生震动的地区的估计”带来一些不确定性。这些更大级别的地震比到目前为止经历的任何地震级别要大，而且是即将要发生的地震。

在设计一个像核电站这样能够经受住地震震动的建筑时，对这个建筑可能经受的地震级别进行仔细评估是非常重要的。反过来，这需要确定可能影响该建筑地点的最大地震的位置和范围。此事件，被认为是该地区有可能经历的“最大可能地震”，从精确度上说，要比该地区已经经历的任何地震级别更大些。估计最大可能地震需要在迄今为止我们已经观察的范围之外进行考虑，即超出我们过去经历之外。这个过程被称为*外推法*。估计最大可能地震需要大量的资金，较大的估计意味着更坚固和更昂贵的建筑，较小的估计则意味着更少的花费。从这个观点来看，与测量仪器的设计和校准、加速度计相关的不确定性就显得无关紧要了。这些不确定性与以下两种不确定性相比显得不重要：一、通过读取安装在这个地区的加速度计的读数从而确定遭受严重震动的地区的不确定性；二、将那些观察得到的数据外推来预测多大的区域将会遭遇更大但还未经历的地震的不确定性。反过来，估计那个地区可能遭受的最大可能地震的范围的不确定性又增加了所有那些不确定性。显然，当把与加速度计读数的准确度和精确度相联系的不确定性，置于从那些读数推导的重要的量

所具有的总体不确定性的背景中时，这种不确定性就很小了。

然而，所有这些不确定性并不能妨碍人们做出决定。行政官员决定在哪里建造新学校、高速公路天桥和发电厂以及如何建造它们。当时的传统工程知识，和经济的、人口统计学的、政治的因素一起，共同指导着建筑标准。考虑到不确定性，这些建筑标准都有一个有益的减震装置。随着经验的累积，从成功设计中学到的经验和失败设计中吸取的教训，会导致对建筑标准进行修订。我们从错误中汲取教训，从成功中总结经验。实际上，标准并不是固定不变的。

因此，现在我们已经看到测量的复杂性了。我们已经考虑了一个特别的测量工具对我们从事的测量的合适性；也考虑到了集合许多单次的测量来判断更大和更重要的变化的方式。下一章讨论的是运用统计学和概率工具对测量数据集合进行定量描述。

注　释：

1. *New York Times*, 22 November 2000, p. A31.

2. Fleischmann, M., Pons, S., and Hawkins, M., *Journal of Electroanalytical Chemistry*, vol. 261, p. 301, 1989.

3. 有趣的是，单词“几何学”本身字面的意思是“地球测量”。

4. 参见 Daniel Kleppner in *Physics Today*, March 2001, pp. 11～12 中的讨论。

5. Dava Sobel, *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*, Walker and Company, New York, 184 pp., 1995.

6. *New York Times*, 11 May and 12 November 2001.

7. *New York Times*, 9 January 1998.

8. *New York Times*, 12 January and 14 February 2001.

第六章　对半的机会

这个国家对信息非常渴望；凡是有着统计特性的每样东西，甚至是一个统计的外观，人们都会对其怀有一种近乎可悲的热情；团体对这些统计性的陈述没有半点怀疑和全面评价。

——弗朗西斯·A·沃克(Francis A. Walker)，

1870年美国人口普查总监

前一章集中讨论了测量的问题。在此我们将更多地谈论如何从所收集的测量数据中提取定量信息。这个讨论将会带领我们进入诸如洪水和地震发生的频率、民意测验及人口普查等多个方面的主题。只有少数投票人参与的民意测验对于即将举行的选举的可能结果能告诉我们什么呢？我们从沿河流域爆发洪水的历史能知道些什么，这些历史对于将来可能发生的事情会给出什么指示？每一个主题都包含着不确定性，来源于不同资料的不确定性，其量化的方法也不同。

我们有很多方法和途径建立测量数据集。一个非常常见的数据来源是对单个物体进行多次测量——当地小学一个班的每位同学测出他们老师的身高，或者是一个地区所有地震观测站测出前一天地震的强度。第二种常见的数据集包含了同一时间内对不同物体测量的数据——或者是新学期第一天上午测量班上每位同学的体重，或者是一个县每口水井在某一天砷的含量。第三种数据集是一段时间内的一组测量资料——一个学校很长一段时间以来在新学年初测得的 8 岁孩子的平均体重，或者是过去 75 年中，当地河流每天的水位。我们如何来定量地描述这些数据集的性质呢？从每组数据集中我们能学到什么呢？

统计学和概率

统计学是定量描述数据集并用已定义的数学程序进行计算的一门科学。学生们把他们各自测得的老师身高集中起来，发现这些值都在 5 尺 9 英寸(1.75 米)左右，有些稍微小点，有些稍微大点。在这里有价值的统计项仅仅是所有测得数据的平均值：所有数据相加除以数据的个数。这个平均值就是老师身高的最佳估计值，平均值两侧测量值的分布情况确定了数据不确定的程度。正是由于学生使用和读取测量仪器的技能不同，造成了数据的不确定性。

第二种数据集，班上学生的体重，呈现出一个 53～81 磅的范围，由此得到的平均体重是 66 磅(30 千克)。在这种情况下，之所以出现一个数据范围，小部分原因是读数不准确造成的，主要原因是因为班上学生的体重不同。我们从这些测量数值得到的信息比仅从平均值得到的信息要多。我们可以确定每个学生的体重偏离平均体重的程度，也可以计算出平均偏离值大小。当我们把学生体重同平均

体重相比时，我们能够知道学生体重分布的密集程度或分散程度。由此可以知道我们班是学生体重几乎相同的单调班级，还是体重超重和厌食学生很多的高度变化的班级。

同样，我们也可以确定在平均体重值周围包含中间2/3学生的体重范围，或者找到一个值，使得一半学生的体重超过该值，一半学生的体重轻于该值。以上所有这些都是对体重测量集合的定量描述。这些测量及其定量描述可以作为对这个群体以及其他地区8岁孩子进行科学研究的起点。这个学校的学生在多大程度上代表了这个县其他学校的学生呢？我们能把这一平均体重或体重范围与这个县的社会经济学因素联系在一起吗？

概率就是可能性估计。对于8岁孩子的情况，我们也许想知道学生的体重介于60磅和70磅(27～32千克)之间的可能性。分析显示学生的体重落在这一范围的概率是75%。同样的计算结果也说明了没有落在这一范围的概率是25%。在本例中，任意选定一个学生，他的体重落在这个范围之外的概率是1/4，所以说存在不确定性。概率为我们定量不确定性提供了一种工具。

很多科学家用下表中的语言描述与不确定性相关的概率的定量估计。有时不确定性用黑色或白色表示，但通常情况下是用灰色。或者如罗素(Bertrand Russell)* 所说，“当一个人承认没有什么是确定的时候，他必须增加一条，有的事情比别的事情更加确定。”以该表作为指导，一个人赢得抽彩的概率我们可以称之为“极度不可能”；太阳明天升起的概率是“几乎确定”。你孩子上大学的概率

* 罗素(1872～1970)，英国哲学家，同时又是著名的数学家、散文作家和社会活动家。他是20世纪分析哲学最具影响力的人物之一，曾因《婚姻与道德》获1950年诺贝尔文学奖。在哲学上，他的观点大抵早期是属于新实在主义的，晚年逐渐转向逻辑实证主义。——译者

则落在极端之间的灰色区域。

概率语言

概率(%)	含　　义
<1	极度不可能
1～10	很少可能或者非常不可能
10～33	有一些可能或不太可能
33～66	中等可能
66～90	很可能或很有希望
90～99	非常有可能或非常有希望
>99	几乎确定

气候变化政府间专门委员会(IPCC)是大约20年前由联合国和世界气象组织共同创建的一个组织。建立IPCC的目的是评估对气候的了解情况和影响气候的因素，评估全球性和地区性气候改变造成的后果，提供未来发展的一系列场景。这些场景可能是整个21世纪人口统计学和经济学发展的结果。在对全球气候各个方面所进行的评估中，IPCC的科学家尤其关注的是与量化观察资料和预测将来相联系的不确定性。把20世纪大气中二氧化碳含量的增加这一观测结果划到"几乎确定"级别。北部中纬度地区由于炎热的夏季导致土壤湿度降低被评估为"很可能"事件。像全球气候系统这样复杂的系统中的每一个要素、每一个过程都有不同程度的不确定性，有些大家很熟悉，有些大家则并不怎么了解。

人们对日常生活事件发生的可能性进行估计的愿望很强烈。明天下雨的可能性有多大？在下一届英国高尔夫球公开赛中，得分低于270杆的高尔夫球手夺冠的可能性是多少？在外科手术中病人存活的可能性有多大？对未来事件不确定性的量化很大程度上是来自于对过去这些事件的统计学分析。对所有以前英国高尔夫球公开赛中获胜者得分的分析让我们很容易在一开始就估计出低于270杆获胜

的可能性。当医生告诉病人“你在手术中存活的概率是50%”时，医生很可能是根据以前做这项手术的其他病人的历史记录总结出来的。一方面，如果这项手术做了上千次，并且在手术中有一半的病人死亡，那么50%存活的概率是从相当多的手术经历中得出的粗略估计。另一方面，如果这项手术只做了两次，并且有一位病人死亡，那么，从过去经历算出的概率仍是50%。然而对于这一死亡率的估计能否代表以后100例手术的结果，对此我们的把握很低。

“从经验中学习”这条谚语通常情况下是正确的。统计学有助于我们估量我们已学到了多少东西。如果一项手术在几十年里做了数千次，开始时死亡率很高，但是随着时间的推移，死亡率急剧下降，那么综合死亡率并不能真正说明下一位病人将要面对的是什么。根据经验，我们会更多地了解到事情是如何运行的。我们希望做手术的医生能提高他们的技术，获得更好的医疗工具，设计避免某些危险的方法，这样随着经验的积累，手术中存活的概率也会改变。

我们把日常的天气预报看成是过程随着时间推移而改进的另一个例子。当我们看电视上的天气预报时，听到明天降雪概率为90%，因此我们可以很自信地预计到明天地面将会白雪皑皑。但情况并不都是这样。几十年前，天气预报是喜剧演员讽刺、嘲笑的对象。公平地说，天气预报的发展已经完全超过了纯粹的统计，比如说以前预报明天有雨是根据过去的128年中有65%的年头在那一天下雨了。今天我们对预报的准确度更有把握，因为我们有卫星摄像机，它能给出整个大陆范围的天气系统图像，气象站组成的全球网络把测量的温度和大气压传送给大型计算机，大型计算机一小时一小时地做出可靠的预测。简言之，预报概率之所以得到改善是由于我们有了更好的

设备，它能给日益改进的计算机模型及时提供更多信息，计算机对信息进行处理得出高准确度的预报。

在外科手术和天气预报的例子中，随着时间的推移，成功的概率都大大提高了。然而，上述事实不能成为推迟手术和不做天气预报的理由。我们清楚地意识到，尽管未来的知识能更好地服务我们，但是我们不能奢望能够等待未来新的知识。不完善的知识总比没有好。

对概率的误解

不幸的是，概率及概率计算的基础经常会被误解。我想起一个老笑话，是有关一个病人面对危险手术的故事。他的手术医生告诉他，他在手术中存活只有对半的机会。病人听到这一消息沉思了一会，然后告诉医生他会继续做手术，但是想把手术安排在有同样病症的其他人在手术中死亡以后立刻进行。他的逻辑是什么呢？如果两个病人手术，有一个病人可能会存活，他想成为均衡前面死亡的那个人。很容易看出他逻辑上的错误：他在手术中存活（或死亡）跟之前任何一个病人的命运绝对没有任何关系。后面病人的命运也不会依赖于他手术的结果。大部分的概率理论，认为每一事件都是独立事件，并不以另外一个事件为条件。在投掷硬币（coin flips）事件里，下一次投掷的结果并不依赖于已经投掷的正面朝上和反面朝上的次数。事件结果的独立性通常是一个容易让人误解或者不能理解的概念。概率也易于被人误解或曲解，尤其是当使用不熟悉的定量语言对其进行计算时。举例说明，让我们看一个潜血检测试纸检验（hemoccult）的例子，简单来说就是检查粪便中是否有血，它是可能患有结肠和/或直肠癌（结肠直肠癌）的征兆。[1]结肠直肠癌发生的统计

结果表明美国的普通人口中结肠直肠癌发生的频率是每 100 000 个人中大约是 300 例，患有结肠直肠癌的人中有大约一半人的潜血检测试纸检验呈阳性(例如粪便中有血)。如果一个人没有患结肠直肠癌，他或她仍有 3%的概率检验呈阳性，所以说，一个人即使没有得结肠直肠癌，他在潜血检测试纸检验中也有可能呈阳性。这只是说明除了结肠直肠癌外，还有其他原因导致粪便中有血。

内科医生如果看到潜血检测试纸检验呈阳性时，他必须确定病人实际得结肠直肠癌的可能性。这个问题没有听起来那么难。300 个患有结肠直肠癌的人中，大约有一半或者说 150 个人潜血检测试纸检验呈阳性。99 700 个没有患有结肠直肠癌的人中，有 3%或者说约 3000 个人潜血检测试纸检验也呈阳性。而潜血检测试纸检验呈阳性的 3150 个人中有 150 个人确实得了结肠直肠癌。因此一个人潜血检测试纸检验呈阳性而实际患有结肠直肠癌的概率是 150/3150，或者说大约每 20 个检验呈阳性的人中就有一个可能患结肠直肠癌。但是已经知道患有结肠直肠癌的人中有一半人的潜血检测试纸检验呈阳性，因此潜血检测试纸检验是医生在病人潜血检测试纸检验呈阳性后，对他进行其他检查的一个信号，例如进行结肠镜检查，就是用来确定阳性的潜血检测试纸检验结果是由于癌造成的，或仅仅属于 20 个人中 19 个检验呈阳性的结果并不表示患有癌症之一。

如果一般的病人对医疗概率感到不可理解，我们或许不会感到特别惊奇，然而，当 1/3 以上的医生在被告知以上事实后，让他们估计阳性的潜血检测试纸检验结果代表患有结肠直肠癌的概率是多少时，他们得不出 1/20 这样的结果，这才更加令人不安。概率性的语言(尽管这些概念并非必不可少)是理解的障碍，而误解通常会导致不确定性。

百年一遇的洪水

让我们来看另外一个例子。当新闻记者评论道“那是在不到10年的时间内爆发的第二次‘百年一遇的洪水’”，很多听众想知道这怎么可能呢？怎么会在爆发第一次百年一遇的洪水后不到100年，再次爆发呢？术语“XX年一遇的洪水”的确表示要经历同样大小的洪水平均要再等XX年。也就是说，平均在XX年内只会爆发一次那么大的洪水。当第二次爆发洪水的时间比XX年早，则说明与平均时间间隔相比，它来得早了些。当然，两次同样大小的洪水爆发的时间间隔比XX年长，则说明时间间隔比平均的时间间隔要大。这种描述性术语不能被解释成排除了洪水多次发生的可能，也不能保证洪水一定在指定的时间间隔内发生。术语注重的只是该流域的平均行为。

水文工作者勉强接受用可能性而不是逆程周期来解释这些概念。百年一遇的洪水也可以被描述为在一年中爆发洪水的可能性是1%。这种描述有助于没有专业知识的读者理解为什么在一个世纪之内会发生一次以上百年一遇的洪水。在任何给定的一年中，都有1%的概率发生这样的洪水。不管什么时候发生，它都是一个不可能事件。

对于给定流域爆发洪水的概率估计，部分是根据年度基础上洪水会涨到多高这些历史性观察资料得出的。在绝大多数年份中，水位发生季节性涨落，但是都会保持在溪流或河道的堤岸内。有时洪水会溢出堤岸，淹没一些周边的地区。它们很少会覆盖整个山谷或者淹没整个平原，很少上涨到山谷的侧面。最常见的是水位波动，根据一个世纪或更长时间的资料记载，如果每4年中有3年会达到一个水位，我们可以有把握地说在一年中水位上涨到那个水位的概率是75%。如果在过去的一个世纪中水溢出堤岸有12次，我们估计在某一年中发生

水位溢出堤岸的概率是12%，对此结论我们的把握要低些。

1924年巨大的洪水淹没了那个假设小村庄皮但克(Podunk)的渡口，1976年又发生了一次，对于这样的事件它的概率是多少呢？我们可以这样说，本世纪发生了两次洪水事件，或者说相隔52年发生了两次大小相当的洪水，但是如果说由此得到的一个洪水爆发的时间间隔就是洪水爆发的平均时间，我们一定会觉得这样的说法非常不可靠。洪水上涨，溢出了堤坝，直淹至房屋二楼的窗户，这样灾难性的洪水发生的概率又会怎样呢？皮但克渡口最初建于1887年，在它建立的127年中从没有发生这样的洪水。但是它会成为这类洪水事件的牺牲品吗？如果会，那么会以什么样的频率发生？对于这类问题的答案，是由建立在较小规模洪水发生频率基础上的概率分布估计和洪水大小与发生频率之间关系的理论推出来的。对皮但克河流域的情况分析表明，这样的洪水相当于一次"两百年一遇的洪水"，这就意味着，平均来讲，这样的洪水两百年才可能发生一次，或者说某一年中爆发这样大的洪水的概率只有1/200。

对于估计一定强度地震发生的概率也可以进行类似的分析。尽管地震学家没有提到"百年一遇的地震"，但概念是一样的。要计算一所马上要建立的学校或核电站可能会经受的最大的地表加速运动，涉及对这些建筑结构寿命期袭击该地区的最大强度的地震进行估计。这样的估计几乎都是根据该地区不同强度的小地震发生的频率推断该地区还未曾经历的大强度地震。如果预期结构的寿命是50年，那么如果结构设计仅仅能抵抗"50年一遇的地震"，这样是不明智的，因为"百年一遇的地震"可能会在不到100年的时间内再次发生。

上面估算的主要假定条件之一是支配洪水和地震发生的基本物理

过程在很长一段时间内保持不变。就地震来讲，这种假设条件是可靠的，因为涉及的现象——地壳构造板块的缓慢漂移，是受到地球内部深层次在地质时间基础上经历变化的过程支配。然而，如果认为某流域的洪水泛滥是保持一个静态的过程，是意味着“导致洪水发生的条件在收集到的流域观测资料的时间间隔内总是保持不变”的技术术语，这就太天真了。

随着一个地区的建立和发展，影响洪水的很多地区性因素也在发生变化。在定义洪水的方程中有一个因素是暴雨中的雨水或融化的雪水渗入地表有多少，有多少从地表流掉，最终汇入这个地区主要的水道。为了发展农业而采伐树木，为街道、停车场铺设路面，这些代表着发展的人类活动都减少了水分向地表渗入，增加了水分的流失。远在上游的土地使用变化也增加了下游洪水泛滥的可能。所以即使几十年来当地的气候保持不变，但是由于城镇化的发展，洪水泛滥的概率也会慢慢增加。

在整个全球气候系统的变化中，年均降水量也正慢慢发生变化，这为洪水发生概率的不确定性又增加了一种因素。影响水分渗入/流出地表的另一个因素是降雨持续的时间。慢而稳的降雨会导致水分更多地渗入地表，片刻的倾盆大雨会导致水分大量流失。记录气候变化的一个明显证明[2]是暴雨发生的频率增加。24 小时内降雨量为 2 英寸(5 厘米)或更多我们就称之为暴雨。90 多年以来，美国经历暴雨的区域也在缓慢增加。

所有这些因素告诉我们，洪水爆发的概率就像与手术相关的存活率一样不是固定不变的。我们不仅要努力根据历史事件确定洪水发生的概率，而且应该意识到由于我们局部地区土地开发使用和全球人类的活动改变了气候，从而导致各种条件的变化使得这个概率也是变

化的。将房屋建造在泛滥平原的人、管理这些建筑的分区编码机构，以及提供洪灾保险的保险公司，都必须认清随着时间推移正在变化的风险模式。

由样本做估计

现在很多国家在选举中流行在选举前进行民意测验，以确定候选人的竞选实力。提到的问题对于参加过此类民意测验的任何人都很熟悉："如果今天选举第七区众议院的代表，你会投琼斯(Mary Jones)一票还是史密斯(Alice Smith)一票？"此后不久，我们会在晚间新闻中听到"600个投票者参与的民意测验显示有42%的人支持琼斯，44%的人支持史密斯，14%的人仍不确定。这些数字的误差幅度(the margin of error)是±4%，所以从统计学的角度来看，这两个候选人仍不分胜负"。

上面这些数据意味着什么呢？误差幅度说明了每位候选人目前在民意测验中的优势评估的不确定性。之所以会产生误差幅度是因为在这次的"抽样选举"中只记录了600个人的投票，然而在几周以后真正进行的选举中，可能会有超过200 000人参与投票。抽样选举的结果也可以这样报道，"600个投票者的民意测验说明，如果选举今天举行，那么琼斯的得票数落在38%～46%的范围和史密斯的得票数落在40%～48%的范围的概率将为95%。由于两个人的得票估计范围存在很大的重叠，因此最终的结果不确定；所以如果选举今天举行，由于两个候选人的支持率非常接近，因而无法有把握地估计谁会当选。"

"候选人X的得票数落在38%～46%的范围的概率为95%"这一论述告诉我们，如果民意测验专家重复这个试验100次，每次都选择600个投票者来确定他们选谁，那么，候选人X的得票数落在38%～

46%的范围有95次。只有5次在这个范围之外。换句话说，民意测验的结果不能正确反映对候选人X立场的概率是5%。如果要降低不确定性，也就是说将概率从95%提高到99%，并且民意测验的得票率落在相同的范围之内，那么就应当有两倍以上的人参加民意测验。显然，要降低民意测验的不确定性，就要增加成本。如果不增加参加民意测验的人数，人们也可以估计出一个百分比范围，民意测验专家对候选人X的得票率会落在这个范围将有99%的把握。当然，这个百分比范围要比有95%把握确定的百分比范围要大得多。很明显，在包含候选人可能得票数的范围和计票落在该范围的概率之间，得有一个折衷(tradeoff)。范围越大，概率就越大；相反，如果范围相对较窄，那么得票率落在该范围的概率就相应越低。一个简单的类比就是向圆靶投掷飞镖，射中靶上任何位置一个较大面积的概率当然要比射中靶心一个小面积的概率要大得多。

绝大多数民意测验专家发现，在参加民意测验的人数和民意测验结果准确的概率之间的折衷，应放在95%的概率水平上。民意测验专家对以下事情有95%的把握：大约600个人参与的抽样试验，那么对候选人竞选结果估计的不确定性范围大约为±4%。参加民意测验的人数越少，成本越低，但是误差幅度越大。一个更大的样本会降低不确定性的范围，但同时会增加成本。而且，大样本的改善非常慢。为了把不确定性的范围减半，就必须有4倍的人参加民意测验。并且人们必须记住，民意测验除了避免成本过高，还必须及时。依赖不断增加参加民意测验的人数这样的想法并不能成功，一方面会增加成本，另一方面误差幅度也只是稍有降低。

但是算术学就足够了吗？不确定性范围的有效估算、民意测验专家报告的误差幅度真的是候选人在民意测验中得票结果的不确定性

的理想评估吗？ 未必。 答案关键在于民意测验专家选择参与民意测验的人的能力。 不管什么时候，当想通过测量从一个大群体中抽取少量样本的特性来估计大群体的特性时，估计结果的可信度依赖于小群体实际上在多大程度上能代表一个大的群体。 简言之，恰当的样本选择是估计可靠的关键。

从电话本中随机抽取的一个名字样本中就包含了很多错误的信息：大约有一半的人甚至没有参与投票，有些人是因为年龄小，有些人是因为未注册，有些人是因为不是当地市民，有些人是因为已经搬走了，有些人是因为认为政治和选举与他们的日常生活无关。 另一方面好的样本选择应当符合人口统计学，样本应该在地理、经济、政治和文化等方面具有代表性。 如果 A 县的合法选民是 B 县的两倍，但是 A 县和 B 县平均投票人数分别约为 47%和 55%，那么应据此来指导民意测验样本的选择。 民意测验专家知道通过恰当地选择样本或增加人数，或者同时采取以上两种措施都可以提高民意测验结果的质量。 改进样本会增加成本，因此权衡民意测验的准确度和达到这样的准确度所花费的成本就不可避免了。

似乎对精心挑选的几千人的投票样本能否成功地预测数百万人参加投票的全国大选的结果总是存在一些疑虑。 就这一点而言，概率理论说得很明白。 被抽样的人占人口多少并不是重要问题；真正重要的是抽取样本的实际大小，即被询问的人的数量，它决定着由抽取的样本确定的概率，代表着整个人口的选举情况。 不管样本是从 10 万人口中抽取的还是从 1 亿人口中抽取的，由少数几千人的样本估计出大群体特性的准确程度令人吃惊，其误差幅度只有百分之几。

我曾在儿子身上验证这个命题，他当时正对收集钱币感兴趣。我们提出的问题如下：我们能根据我们家饼干罐中堆积的硬币估计出

最近10年内每年铸造的硬币数量吗？铸造的新币由于流通的需要产量年年都有变化，但是流通数量通常上亿。一个装满几千个硬币的饼干罐能够作为每年铸造的上亿硬币的代表性样本吗？的确可以。饼干罐中收集的硬币比估计铸造硬币数量的年度产量变化所需硬币要多得多，并且误差幅度很小。概率理论有助于我们决定为了成功得到误差幅度范围内的结果所需要的样本大小，盛满硬币的饼干罐对得出正确结果大有帮助。

追踪一段时间内的选举民情的民意测验反映出随着选举展开选民观点的变化。先前的观点可能改变了，“未做决定者”最终也做出了决定。就像生活的大多数方面一样，选举不是静止不变的。许多因素能在选举上主宰沉浮，包括候选人为了发布消息所筹措的资金、候选人的过失行为和错误言论以及媒体对待候选人的态度。最终的结果是运动的目标，候选人尤其关注民意测验中民意的变化和对引起民意变化的分析。投票人做决定的过程是一个非常灵活的过程，竞选的结果很少和最初民意支持结果一致。

人口普查

在上一章中，我用美国10年的人口普查来说明计数的不确定性。除了普查过程中人的生老病死使得“正确的计数”难以获得外，到2000年底在美国居住的实际人口数最终也不确定，对此任何人都不会感到吃惊。在人口普查中，登记人口数就像在选举中清点选票数一样，最终清点的准确程度会受到限制。然而，由于人口的地理分布是划分立法区域和选举区域的重要因素，因此在那些重要的州及国家立法团体中，不同人口统计群体的相对代表中所发生的政治权力的重大变化都是根据不完善、不准确的人口普查得到的。

根据宪法规定，每 10 年对人口数量统计一次。可是每次统计人口都会发生关于如何改进全国人口普查的争论。一种观点是我们必须努力坚持通过使用改进的、更深入的直接交互方法(邮件、电话、因特网和现场会见)数到每一个人。然而，相反观点支持采用抽样策略估算在直接普查中漏掉的人的数量。例如，如果邮寄的调查表和人口调查员只能到达邻近地区 88%的家庭，对剩下的 12%家庭要根据用传统的直接方法数到的 88%家庭的特征性反应进行估计。人口调查局认为，用从来自全国的 314 000 个家庭得到的调查数据计数，可以可靠地估计多计和少计。这些家庭由遍布美国的几百万个家庭抽取的样本组成，人口调查局称之为 ACE 群体，ACE 是准确性(accuracy)、覆盖度(coverage)和评估(evaluation)的首字母缩写。

抽样技术已经广泛地应用到与人口普查相近的更多领域。产品的销售能够通过估算家庭的年收入、拥有的汽车数量、住房的平均价格以及从政府人口普查数据中得到的类似经济特征大大受益。那些出售商品给美国消费者的人，对用不完全的人口普查得到的数据来估计整个地区的人口状况和经济状况充满信心。

此外，对抽样科学理解相对比较透彻，没有什么争议。同时运用简单直接计数方法和估计多计、少计的抽样方法能够成功得到更准确的普查结果，对此也很少有疑问。然而，用来填补计数和计数覆盖范围之间间隙的统计学方法却受到了强大的政治性抵制。之所以有这种政治抵制，并不是因为反对统计学方法的政治家不相信抽样方法。实际上，在大选之前的民意测验中，代表政党的候选人通常很放心地把抽样方法作为反映当前投票者民意的指示器。在人口普查中，反对使用统计方法是因为只采用直接的计数方法得到的不完全普查结果会对某些政党有利。保持不同地区和不同种族的人口普查在

完整性和精确度方面不一致，对从中获益的政党来说是一个不言自明的特殊政治目标。通常，处于不利地位的包括少数民族、移民和穷人，他们相对于那些社会经济条件好的群体来说更容易被过低估计。对这些群体的过低估计会降低这些群体在政府立法机关和美国国会中的代表性。

在本章，我们大概了解了测量数据和观测资料是如何通过统计学方法进行定量描述的，以及如何估计包含在数据集中的信息的质量和可靠性。这是一个内容丰富的话题，关于这一话题吸引了几乎数以千计的短文、书本和专著。显然，我们只是打开了通向这片巨大知识领域的大门的一条缝。但是，通过这窄窄的一瞥，只要人们能够认识到无论是科学还是建立在有限观察资料定量分析基础之上的任何努力，确定性通常是无法完全达到的，这就足够了。有限的观察资料集通常定义了一个解释的范围，并且估计出真实性落在该范围的概率。一个较大的概率范围，也可以说是一个较大的目标，真实性落在该范围的概率也较大，但是由于缺少特异性，也许没什么用处。我们可以描绘出一个较窄的解释范围，但是发现真实性的概率就会相应降低。本章我们学到了从不完整数据能够收集多少信息，这为下一章我们讲述如何概念化和模拟复杂系统提供了一个有用的背景。

注　释：

1. 这个例子由 Hoffrage, U., Lindsey, S., Hertwig, R., and Gigerenzer, G. 在下文中作了叙述，Communicating Statistical Information. *Science* vol. 290, pp. 2261～2262, 2000。

2. Karl, T. R., Knight, R. W., and Plummer, N., Trends in high-frequency climate variability in the twentieth century. *Nature* vol. 377, pp. 217～220, 1995.

第七章 我不太确定
这是如何工作的……

困惑是获得知识的开始。

——卡里尔·纪伯伦(Kahlil Gibran)

统计分析中测量的量化，温度随着时间变化趋势的发现，或者针对一幅图中显示的水污染数据模式的观察——所有这一切都激发着科学家开始思考他们正在观察的关系背后存在着一种什么样的过程，并将这种思考明确地表达出来。最初这些思想是简单和基本的，到后来也许需要通过试验进行检验。在下面两章中，我们将探究概念和试验的世界，了解不确定性是如何促进创造性发展的。

科学家，实际上我们每一个人，总是通过事物运行方式的简单化概念进行工作的。我们将这些简化的表示称为“模型”(models)，它们以多种方式出现：概念模型、物理模型、数值模型。在从现实世界建立的模型中，通过对我们努力理解的现象或系统的不完全、有时是不准确、偶尔是冲突的测量或观察，我们接受到的是不完全的指

引。在模型和观察间有一个持续的相互作用，并且一个在另一个面前经历着调整。新的观察导致一个概念的修改，反过来一个新的概念，暗示需要进行新的试验或观察，以便再一次获得检验。通常正是这样前后迭代的相互影响，才会增进对系统的理解，并且在一些环境下能够减少与系统行为相关的不确定性。但是当这种流动性消失时，例如当一个科学家在相当多的反面证据面前，支持或坚持一个概念或者过于相信不准确的或不相关的观察时，这种过程就停止了。

自然界中的绝大部分，其结构和组织都有着令人难以置信的复杂性。让我们思考以下森林生态系统的复杂性吧，在那里，树木、真菌、微生物、鸟类、小型啮齿动物、大型哺乳动物、昆虫、蕨类植物、蜗牛、青蛙、蛇和许多更多类型的生命，在森林限制的范围内，是相互依存的。它们不但共存，而且相互依赖，每一个都为另一个提供得以繁荣的一些要素。细菌存在于土壤中，影响土壤的化学组成，促进它的发展；微生物也存在于大型动物的消化道中帮助它们进行消化和新陈代谢。而且所有的物体，不论大小，都受其所在区域的天气和气候所影响。由于这种复杂性，即使是最有能力的生态学家，要研究森林生态系统的全部细节，也是十分困难的事情。因此他们发展了有关生态系统工作的简化概念，集中在一些被认为是特别重要的组元以及这些组元的相互作用。这种相互作用的生态网络的概念化(conceptualization)被称为一个模型。自然，不同的生态学家对相互作用的认识是不同的，他们对不同组元的参与程度的权衡也是不同的，从而也就开发出了不同的模型。正是因为以上所述的复杂性，人们对生态系统的理解是不全面的，关于生态系统是如何被整合到一起的不确定性也就随之产生了。

同样，一个工业大国的经济，是在制造商、交通系统、批发商、

零售商、银行、证券和商品市场、顾客、农民、劳动力、政府、税法以及更多的参与者之间的相互作用中形成的复杂网络。对于一个森林生态系统而言，以无限精细的方式描述这个系统，实际上是不可能的，因此简化和综合就发生了。进口、出口集中在一起作为“贸易平衡”；生产能力、工资水平、税、利率、商品的丰富性等等，放在一起形成一个“重要经济指标”的指数；不同商品和服务的价格聚集成“消费价格指数”；数百万人在花费、储蓄、投资或退休方面做针对个人合算的决定时的态度被表述为“消费信心指数”。经济行为，作为一个整体，被认为是由经济学家神化了的一些数量关系的结果，这些数量关系与贸易、消费价格指数、消费信心指数和其他一些聚集指数的平衡相联系着。被聚集的组元数量和组元之间的关系，构成一个经济模型。之所以会产生不同的经济模型，是因为不同经济学家对于经济组元的关联性和交互性所做的假定不同。在美国，有许多模型需要选择和评估：国会预算办公模型、沃顿商学院模型、参议院财政机构模型、密歇根大学模型。每一个模型都加入了它的创立者的判断和观点，而且将会证明一些模型必然比其他模型更具洞察力和先见之明。

概念模型

也许最简单的模型是概念模型。一个概念模型是关于一个系统以及它的组元和组元相互作用的智力想象图。它是更多精制模型，如物理模型或数值模型的基础。一个概念模型提供了一个框架，在该框架中思考系统的运行或者通常要解决的问题。接下来遇到的操作模型(operational model)并不比概念化更重要。

经常在经济概论课程中出现的很熟悉的一个概念模型是供求定

律。供求的概念联系着价格机制中的生产和消费。如果某种产品或商品产量丰富，当前价格超过了消费者对这一产品的需求，那么销售商很可能会选择降低价格来使商品更具吸引力，否则生产者可能会因为该产品不是卖得很好而选择减产。相反地，如果一种产品按照当前的价格卖得很好，商人不可能把它按原价搁在货架上，而会提高价格以便利用这种旺盛的需求。同样，产品制造者会因为该产品大受欢迎而选择增加产量。或者一个竞争者，看到开设一条新的生产线的机会，也会决定开始生产。在一个理想的自由市场中，生产、消费和价格不断进行调整以维持这些因素间的平衡。供给和需求，作为一个简单的概念模型，在广阔和综合的术语范围内，已经相当好地描述了自由市场经济的运作。

然而，如果一个概念模型在它的想象中有太多的限制，或者加入了错误的假设，那么它预期一个系统行为的能力将被限制和(或)误导。在一个理想的自由市场中，供给和需求也许运行良好，有许多的供应商和消费者。但是当供应商存在垄断、销售渠道不足，或者价格受到调控机构的控制或补贴政策的支持时，市场相互作用就会弱化很多。在欧洲经济共同体内部，由于欧洲国家费尽心思维持自由贸易与确保在其经济中保持很强的农业成分需求的平衡，因而有关农业补贴的争端一直是一个恼人的问题。

几十年来，美国电力工业的特征是所有的组成部分已经背离理想的供求经济。在几个州正在进行或者正在计划的公用工业的取消管制，将会以未被验证的方式变更着正在运作的领域。但是加利福尼亚州刚一开始的经历提供了引人注目的证据：取消管制并不总是按照计划进行。[1]电信业领域的取消管制已经开展了10多年，在发送和接收声音、图像、信息以及数据的方式上发生了极其显著的变化，其中

的许多变化在取消管制期间是没有预料到的。自然，在信息技术方面也有许多令人吃惊的事情有待揭示。

在一个非常基础的层面上，20世纪80年代航空工业的取消管制认为竞争的增加将通过更低廉的价格和更优质、更方便的服务更好地服务于顾客。而对于航空公司而言，取消管制为减少运营成本创建了一个新的框架。这很快导致了中枢辐射式航线系统的建立：每一条航线选择几个以飞机为主要运载工具的中心城市。在某些情形下，市场支配等价于事实上的垄断，这种垄断表现为伴随服务的质量下降和对费用进行削减的刺激减小。当小一些的公司在几个中心城市努力争夺部分市场时，更大的航空公司会暂时降低价格，以便把小的竞争者驱逐出市场。很快，顾客无法求助于其他公司的运输工具，因此对于已经建立的航线来说，几乎不必担心缺少顾客。供求概念模型没有被证明是航空工业适应和利用取消管制的很好描述。

运营成本的最小化也在2001年9月11日造成了损失。航空公司把航空安全视为一个需要管理的“成本中心”，同私营安全公司签订合同。这些私营安全公司转而会以最低的工资雇用安全人员，不给额外补贴，经过最低限度的训练就让他们工作。世界贸易中心悲剧的部分原因就在于不重视航空安全。随后由美国国会和总统做出的由联邦统管航空安全的决定，是对航空公司在提供充分安全方面失败的承认。

不完全的概念化

在科学上，模型可能是好的，也可能是坏的，这取决于科学家对系统的概念化程度如何：他或她在理解一个复杂系统的运行时具有多深刻的洞察力。让我们看看科学史上不完全的或者有缺陷的概念化

的一些实例，以及对这些不充分的概念化的修订是如何进展的。

地质学教科书的作者总是喜欢引用一个广为人知的、19世纪时确定地球年龄的方法，并将其作为一个不完全而且易于误导的概念化的例子。确定地球年龄长期以来一直是地质科学的一个中心主题。今天最可靠的对地球年龄的估计是使用放射性元素衰变的方法，例如铀衰变成铅，是通过各种有规律而众所周知的速率进行的核过程来实现的。然而放射性直到19世纪末才被发现；在此之前，科学家是用其他方法来估计地球年龄的。不完全的概念化和最终的错误估计是由苏格兰物理学家汤姆生（William Thompson）* 做出的，该物理学家更以他的贵族头衔——开尔文勋爵著称于世。

开尔文的推理是沿着这样的线路进行的：地球在其初始时期被赋予了一定的热量，从那以后一直在冷却。如果一个人能确定地球最初具有多少热量，而且对于它如何损失热量有一个好的理解，这能够使他对冷却过程已经持续了多长时间进行估计，或等同于自地球形成以后已经经历了多少时间。开尔文的方法论好比一个注满水的浴缸；在某个时刻将塞子拔出开始排水，过了一些时间，水位开始下降。如果人们知道水通过排水管离开浴缸的速率，他就能够通过降低的水位计算出是什么时间拔出塞子的。

开尔文推理：最初赋予地球热量的上限能够通过岩石熔化的温度来确定，而岩石熔化温度在火山爆发地点能够进行测量。这一逻辑暗示人们，地球最初是在一种熔化的状态中形成的，但这之后很快固

---

* 汤姆生（1824～1907），英国物理学家，1892年被封为开尔文勋爵。在他的研究工作中，以热学和电学及它们的应用等方面最有成就。1848年创立绝对温标（亦称开氏温标）；以后，他把热力学第一定律和热力学第二定律具体应用到热学、电学和弹性现象等方面，对热力学的发展起了一定作用。此外，还制成静电计、镜式电流计、双臂电桥等很多电学仪器。1866年起，他领导完成了横越大西洋海底电缆的安装工作。——译者

化成我们今天所看到的岩石。接着他争辩道地球将通过一种被称为热传导的过程损耗热，这种传热方式发生于诸如组成今日之地壳的岩石之类的坚固物体内。他对于此过程的概念化与当营火熄灭后围绕营火的岩石慢慢冷却的过程没有什么根本不同。在一个家庭环境中，传导是热茶倒入茶杯时使得茶杯变暖的一个热传输过程。

开尔文知道在整个过程中冷却速率不是均匀的；地球在形成后很短的时间内冷却得很迅速；随着时间推移，冷却变得缓慢了。因此，如果他能够确定地球在当前损耗热的速率，他将能够判断地球的冷却已经发生了多长时间。他继续对地下矿物做细致的温度测量，以便确定当前损耗热的速率，接着运用此速率计算地球的冷却已经进行了多长时间。他对地球年龄所作的几千万年的估计，尽管比《圣经》估计的几千年要长很多，但却比地质学家用于解释他们所观察到的地质地形所作的必要估计和生物学家认为对物种进化是必要的时间估计要短很多。因此有了19世纪最激烈的智力争论：一方是以地质学家和达尔文为代表的关于地球非常古老的主张，另一方是物理学家开尔文关于地球相当年轻的观点。

这一争论没有解决，仍在持续着。每一阵营都认为对方错了。直到19世纪末放射性被发现以后，开尔文的主张才得到了阐明。不稳定元素的放射性衰变，是由法国物理学家贝克勒耳(Becquerel)* 于1896年首先观察到和描述的。它是我们在原子能核电站运行中已经很熟悉的一个热源。在核电站，衰变能量被转换成热，用于产生驱动发电的涡轮机所需要的蒸汽。放射性衰变对于地球年龄的争辩有

---

* 贝克勒耳(1852～1908)，物理学家，放射性的发现者。贝克勒耳发现天然放射性，标志着原子核物理学的开始。由此他和居里夫妇共同获得1903年的诺贝尔物理学奖。——译者

着什么样的重要性？开尔文所作辩解的关键是，仅有一个热源损耗热量，地球从它最初熔化状态下获得热量。由于岩石包含少量放射性元素，在地球原始热量损耗的同时它们会提供新的热量。这当然意味着要比没有任何额外热源时冷却要持续更长的时间。回到确定浴缸排水多长时间的类比，那么所做的实验就好像是在考虑通过排水管排水时，没有意识到这是在水龙头是开着的、还在往里面放水的情况下进行的。

开尔文的计算并没有错误，但是他却将这一问题不完全概念化了。他没有办法知道：地球热量的估算不仅应包含继承的，还应考虑从放射性衰变中获得的热量。由于将要被发现的放射性的存在，加上地球岩石中存在着放射性元素，因此开尔文关于地球年龄的确定是很不准确的。极具讽刺性的是，开尔文确定地球年龄的方法中未使用放射性现象，而放射性现象在后来被证明是构成当前确定地球诞生时间的最可靠方法的核心要素，就像前面讲到的，这一方法建立在一个放射性元素（“母体”）到一个稳定的元素（“子体”）的稳定衰变的基础之上。老的岩石与年轻的岩石相比，具有更少的母体和更多的子体。母体铀和它们的子体铅已经显示地球比 45 亿年略微老一点。

有缺陷的概念化

科学史中充满着有缺陷的概念化。其中一些在说明误解是如何阻止理解，结果又如何促成不确定性方面特别有益。首先让我们看看哥白尼(1473～1543)关于我们的行星系统的结构和动力学的著名再构造，这一构造大约是在 1530 年完成，直到他死前不久才被出版。

人类对围绕我们的宇宙的好奇很久远了。3000 多年以前，巴比

伦人就喜欢观察夜空。他们注意到，在我们现在称之为星座的几何图案中，绝大多数的光点保持着相对的固定。但有一些明亮的点跨越这些图案移动，巴比伦的天文学家称这些移动的亮点为“未驯服之羊”，在希腊词语中用“漫游者”来表示，今天我们知道这些未驯服之羊是太空中我们的近邻，我们将其称为行星。希腊人将他们自己的观察加入波斯人和巴比伦人的观察之中，发展出了一幅以地球为宇宙中心的天体图。在这一“地球为中心”的图景中，太阳、月亮和行星环绕地球运动，就像侍女服侍一位君王。作为这个幅员辽阔的王国的居民，星座留在背景之中。后来，这一观点出现在托勒密(Claudius Ptolemy)的著作中，托勒密是公元2世纪亚历山大的天文学家和地理学家，亚历山大是当时知识分子生活的中心，位于现今开罗的一个郊区。托勒密关于宇宙地理学的观点在哲学和宗教上有着巨大的吸引力。这种以地球为中心的宇宙被视为上帝的杰作。上帝创造了地球上的人类作为宇宙的中心物，创造了其他的一切事物作为背景。

托勒密的观点实际上影响了欧洲和地中海各国人的思想，直到这一概念化受到哥白尼的挑战。哥白尼在接受传统的以地球为中心的宇宙观点的教育时，发现这种观点不是令人很满意。他的直觉引导他坚信：正是作为巨大热源和光源的太阳，占据了行星系统的中心，而地球仅仅是围绕太阳转动的另一颗行星。根据这种观点，显著的太阳和月亮的每日运动以及跨越天际的星座的年度移动，根据地球每日绕轴的自转和每年绕太阳的公转很容易得到解释。作为一个附带的好处，从地球上看这些行星的运动的描述与在以地球为中心的系统中的描述相比较要简单多了。哥白尼相信：一个简单的解释比一个复杂的解释更吸引人。这种简约的原则在两个世纪以前就

被奥卡姆的威廉(William of Occam)* 清楚地阐明了，关于奥卡姆，今天为大众所熟知的是“奥卡姆剃刀”**。（奥卡姆，通过表达打破旧习的观点来反对他同时代的当权派。他也发现自己由于背离了传统智慧而不受欢迎。）

但是哥白尼作出多么大的挑战啊！实际上，这种挑战相当于在皇室法庭，把地球从一位崇高的王国君主骤降成依附他人的贱民团队的成员。这是把地球从宇宙中心降格到太阳的第三行星(the third rock)的重新安排，也是对宗教和哲学机构的冒犯。这些机构已经越来越习惯于这样的现象：重要性、权力和财富总是伴随着上帝造物的伟大方案的官方解释者。

然而，思想在最初建立时，无论是否得到广泛的信服，都有着它们自己的生命。以太阳为中心的行星系统概念建立了在其中成长的其他丰富的思想。哥白尼的思想激发了第谷·布拉赫(Tycho Brahe, 1546～1601)对行星运动进行进一步艰辛的观察，开普勒(1571～1630)从第谷的数据中总结出了他的著名的行星运动三大定律的特定规律。牛顿(1643～1727)发展了一个关于力及其相互作用的优美的数学描述，从这个描述中能够推出开普勒的行星围绕太阳运行

---

* 奥卡姆的威廉，中世纪英国哲学家。名叫威廉，出生于英国的奥卡姆，人们称他为“奥卡姆的威廉”。主张唯名论，只承认确实存在的东西，认为那些空洞无物的普通性概念都是无用的累赘。——译者

** 奥卡姆剃刀，是指奥卡姆的威廉的著名格言“如无必要，勿增实体”，他挥舞这个剃刀的本意，是为了说明上帝的存在不能通过理性推导而得。奥卡姆剃刀后来也被当做科学研究和理性思维的一条原则，但在具体应用时有多种变形。其中一种是：不应加入无必要的假设，在两种等价的结论中，应选择简洁的、假设较少的一种。例如有以下两种结论：一、生物经过进化而来。二、上帝创造了进化的准则，生物经过进化而来。这两种结论说明的是同一个事实，即“生物经过进化而来”，但是后者的假设“上帝创造了进化的准则”对说明这一事实没有任何必要，因此是个多余的假设，在科学上应用奥卡姆剃刀将它剔除。奥卡姆剃刀的另一种表述方法是：在多种可能性中，应选择最简单的那种。这并不否认复杂可能性的存在，而只是说，在没有证据时，应该首先接受、试验最简单的一种可能。——译者

的轨道。今天，我们称这一相互作用为万有引力，行星系统就是众所周知的承认太阳中心位置的太阳系。

哥白尼革命——以地心的或以地球为中心的概念化被以日心的或以太阳为中心的框架所替代，其重大意义在于，对一个复杂现象的理解取决于被研究系统的一个适当的概念化。在放弃地球位于太阳系中心的观念之前，对太阳系结构和动力学做到深刻理解是不可能的事情。然而，一旦日心说成功了，误解的桎梏被打破了，接二连三的阔步发展就发生了。今天我们从地球轨道卫星发送的每日的天气云图中受益，我们可以看到月球上的宇航员或一个机器人漫游者从火星上给我们发送的照片，没有人不为之激动。我们也惊骇于弹道导弹轨迹的精确性。所有这些之所以成为可能，是因为关于行星系统的一个适当的概念化实现了理解的进展。

大陆漂移

另一个说明有缺陷的概念化如何阻碍进步的例子，是从地质科学史中选取的。整个地质时期大陆在地球表面漂移的概念最早是从地理学家那里得到拥护的，这些地理学家注意到在非洲西海岸和南美东海岸的外形之间有着巨大的相似性。用富含想象力的眼光来看，这两个大陆就像拼图玩具的碎片一样能够整齐地结合到一起。从这一表面上的相似性，地质学家推断南美和非洲曾经是连在一起的，只是在遥远地质年代的某个时候裂开而分离了。培根(Francis Bacon)爵士在17世纪早期曾对这一非凡的地质现象发表过评论，在18世纪其他人也作过论述。到19世纪末，地质学家已经发现沿着两大洲的边缘在对应点的岩石和化石非常相似，这实际上表明，不仅外形拼接是合适的，而且拼接出的图案也是连续的。20世纪早期，这种地理和

地质学的证据由德国科学家魏格纳(Alfred Wegener)在一本如今非常著名的题为《大陆和海洋的起源》[2]的著作中进行了总结和阐述。魏格纳介绍了一个引人注目的事实：大陆曾经是连在一起的，从那以后漂移分离。

魏格纳的概念，尽管对那些研究南半球地质的人特别有吸引力，但并没有受到那些研究岩石物理特性的另一群地球物理学家的赞赏。地球物理学家指出，在南美和非洲之间，排列了大约3000英里的固态岩石，它们隐藏在大西洋的海水下面，形成位于海洋之下地球的地壳。地球物理学家将大陆漂移概念化为一个类似于船只乘风破浪穿过海洋前进的过程。他们问，大陆如何能够通过这些岩石的障碍前进呢？他们争论说构成海洋底部的岩石实在太坚硬，不允许出现这样的运动。由于缺少大陆漂移的明显机制，因此认为大陆在地球上方移动的地理学和地质学观点被人们忽视了几十年。拼图玩具的合理性被人们看成是类似另一个著名的地理上的古怪性——意大利看起来像一只靴子，而漫不经心地摒弃了。至于大西洋沿岸岩石和化石的相似性，他们能够从相似的过程和平行的进化中得到结果。

这种状态一直持续到20世纪60年代中期，这时出现了一个叫做板块构造学(plate tectonics)的新概念，对大陆漂移的古老争论进行了重新评估。这一新概念承认大陆不能通过海洋底部的地壳岩石前进。相反，这种观点主张大陆和围绕它们的海洋岩石一起移动，很像一根冻在一片移动着的冰层的圆木。当冰层和它们的圆木“大陆”沿着主要的破裂处彼此分开时，在它们之间露出的水会冻结，在空隙处形成了新的“岩石”。

对构成洋底岩石的年龄的重要新观察激励着上述关于地球的修正主义观点。实际上这些数据并不是很新，收集起来主要用于在第二

次世界大战和随后的冷战期间与潜艇战相联系的军事目的。当这些数据公布给科学共同体时，它们揭示出洋底的岩石全都比大多数的大陆岩石要年轻，就像在湖面上分离的冰层之间新形成的冰要比这些冰层本身年轻。因此地质学家所收集的关于大陆漂移的充分证据，不再会由于大陆漂移在机制上的不可行的争论而被搁置一边了。大陆移动的概念已经被抛弃了，原因在于“大陆通过地球刚硬的地壳主动地推动自己”这样的错误概念。一旦意识到大陆是一个骑马旅行的被动乘客，地质学的证据很快会得到有效的承认。

就像哥白尼的革命性思想导致对太阳系的理解取得巨大进展一样，地球科学同样也从新近发展出来的地球动力学的板块构造概念受益非凡。新的海洋学数据使得上述这种关于地球的新观点成为可能。此观点激发了对实际问题的很多创造性思考：如石油、天然气的生成，矿床的形成，地震和火山危害的性质，影响长期气候变化的因素，等等。以前的观点被放弃了，原因在于新观察迫使进行重新思考，新的概念化能够在适应新的观察的同时与更老的证据保持一致。板块构造学远去了，地球科学永远在发展。

墨守成规与标新立异

墨守成规很少被认为是有积极作用的思想。通常我们把这种情形同挫折、不满以及缺乏思考问题的新方式相联系。有人恳求我们“在盒子之外思考”，打破限制我们概念化问题的思维障碍。苏格兰阿伯丁大学的物理学教授琼斯(Reginald V. Jones)，给我讲了许多关于避免概念性惯例的重要性以及尽可能快速识别惯例的价值，以便在惯例陷得太深和时间浪费太多之前能够摆脱出来。

琼斯，第二次世界大战期间英国的一个年轻小伙子，在科学研究

和应用于战争的情报方面扮演了一个重要的角色。他在其《最隐秘的战争》[3]一书中，讲述了一个逃避惯例的故事。他的团队早期的任务之一是研制一种不能被雷达发现的飞机。雷达已经扮演了重要的角色：一是为英军提供“敌人正在逼近并空袭英格兰岛屿”的早期预警，二是警告德军盟军的飞机正在朝欧洲的目标挺进。如果琼斯能够设计一个不让雷达发现飞机的方法，那么盟军将具有明显的优势能够轻而易举到达目标而不受德国空军拦截机的干扰。今天这一概念已经在美国空军“秘密行动”飞机家族中变为现实。

琼斯和他的队员试验了各种技术技巧，都没有效果。一种木制的飞机，机翼和机身都有特殊的几何形状，并且涂胶，仍然不能防止巨大的金属发动机强烈的雷达反射。他们反复试验，都没有成功。琼斯意识到他们在墨守成规，没有在盒子之外思考。最终他想到了一个解决方法：他问道，“你如何隐藏一粒沙子？”回答是直接的：你将它放在沙滩上，融入无限多的沙粒之中。将此运用到飞机伪装业务上，解决方案不是隐藏单个雷达信号，而是创造100万个雷达信号。由此诞生了从诱饵飞机上倾倒碎金属片的战略，这种金属箔片如同我们把食物包裹起来存储在冰箱中所用的。对于敌人的雷达来讲，每一个金属碎片看起来像一个到来的袭击者，而这些碎片在整体上就如同一个由极具破坏力的复仇的轰炸机和战斗机组成的实际编队。防御拦截机起飞来对付这一优势军力，到达指定空域却发现没有飞机。与此同时，一个很小的真正的轰炸机群，沿不同方向飞到敌人的领土上空，实际上可以毫无麻烦地投下高爆炸性炸弹。这一欺骗虽然只持续了几周，但在那一时期会造成相当大的破坏。

然而为了能够“在盒子之外”有效思考，我们首先必须意识到，我们实际上是处在限制我们思考问题方式的“盒子里边”。这些时常

微妙的限制是什么呢？ 让我们以一厢情愿的想法（wishful thinking）开始吧。 当社会事物或经济事务进行得不是很好时，我们经常对自己说“事情会变好的”。 当一个投资者喜爱的股票已经经历一个长时期的痛苦下跌时，在下跌的许多阶段，他或她将要争辩说情况不会变得更糟糕，股票的价格已经降到最低了，很快就会走上恢复的道路。 在这样的情况下，如愿骨支配着，而脊柱弯曲着。 这种态度常常是对许多紧迫的环境问题的反应。 我们认识到空气质量已经恶化，马路和高速公路非常拥挤，海滩的水对于游泳是不安全的……但是我们对自己说情况会好的。 然而，几乎没有任何问题通过自发的补救幸运地得到解决，或者仅仅通过期望就能消失。 只有对问题进行关注、分析和执行之后，才会发生变化。

爱因斯坦奉行这样的信条——“你不能运用当前思考解决当前问题。 当前问题是当前思考的结果。”当前思考仅仅是传统智慧的一个别名。 作为解决问题的障碍，传统智慧表现为一系列不同的形式：我们总是以这种方式做这件事；我们以前试验过；这一纠正花费太多；那条路注定通向失败。 传统智慧是没有经过检验的重复。 这是对当前观点的一种接受，没有问支持这些观点的基础本身是否经历着变化。 拒绝或挑战传统智慧是用令人不安的不确定性来代替明显的确定性的一种行动。 如果哥白尼或者魏格纳只是接受了他们同时代的传统智慧的话，今天我们在理解太阳系和地球如何运转时会怎样呢？

意识形态是一种特殊形式的传统智慧，它通常受到一些重要影响的支持。 政府和宗教机构有时会建立传统智慧，接着保护它们不受常规的挑战。 在这一意识形态的防卫之后，人们经常会发现一厢情愿的想法。 托勒密的以地球为中心的行星体系的概念已经成为哥白

尼之前的意识形态，原因在于当时的宗教机构和非宗教机构都想让它成为现实。一段时间以后，苏联的遗传学落后了，原因在于科学思考被强安了意识形态的盒子。在建立被误导的遗传学的过程中扮演了关键角色的人是李森科(Trofim Lysenko)，此人是在乌克兰农业试验站从事研究的生物学家和农学家。他宣称小麦种子在合适的环境条件下能够被用来生产黑麦，而且这样的能力能够遗传。这对苏联的意识形态特别具有吸引力，苏联的意识形态是：人类在社会性上是可塑的；良好的社会主义观点不仅能够习得，而且能够继承。苏联的当权者想让李森科的农业思想成为现实，一度禁止讲授和研究在世界上其他地方蓬勃发展的遗传学的主流思想，这是对意识形态的忠诚而不是对科学的忠诚。

21 世纪早期，不承认全球性气候变化正在发生以及人类在引起这些变化中扮演重要角色的观点，是以一厢情愿的想法和传统智慧为基础的。化石燃料工业和交通运输业不愿相信一些非凡的事情正在发生，这些事情可能会强迫它们重新构造，甚至更糟糕的是强迫它们放弃已经持续了一个世纪的成功故事。我们已经利用化石燃料一个多世纪了，这件事情意味着“我们一直使用化石燃料，将来也一直使用”。这种观点，只要是流行并具有说服力，将会引导我们进入一个持续加深的墨守成规之中。

一厢情愿的想法很少是一个成功的长期策略，现实迟早会迫使我们承认变化。人类历史上的石器时代之所以结束并不是因为人类从石头中走出来了，而是因为人类懂得通过金属可以制造更好的工具。类似的，人类发展的化石燃料阶段也会在化石燃料耗尽之前结束。在漫长的人类历史中，无论过去还是未来，在人类学会如何利用和集中由太阳直接照射到地球的实际上是无尽的辐射能量之前的时间间隔

内，对化石燃料的依赖只能是作为一个暂时的帮助。实际上，煤和石油之类的化石燃料，是自然界中比较拙劣和低效的产品。这些产品埋在地球地壳中，以化石化的植物和微生物的形式储藏着过去的太阳能，并通过“烹饪”的方式分解和重组这些能量。

物理模型

小时候，我玩林肯圆木，修补玩具和装配套件，构造无限变化的圆木盆、风车和金属梁的桥。一段时间以后，我开始玩飞机模型，煞费苦心地削出一些西印度轻木条做成机翼和机身，然后给它们覆盖一层纸质蒙皮。用一个螺旋桨和橡皮带提供动力，微型飞机就可以飞向天空。所有这些儿时的小玩意都是一些现实世界构建物的简单化描述，是对现实世界实际建筑按比例缩小的物理模型。作为物理模型，它们具有现实世界建筑的许多特性，但是它们也会以其他重要的方式与对应的部分相背离。一个简单的飞机模型用纸而不用纤维或者金属做它的蒙皮。用橡皮带的弹性能量而不用来自液体燃料的化学能提供动力，而且这种模型没有飞行员，也不需要控制。如果一个人试图建造一架模仿这一模型的真实飞机，那么他会很快认识到这一模型的不充分性。事实上，该模型飞机只是简单的飞行空气动力学原理的一个模型，而不是一架在起飞、飞行和着陆期间需要控制的飞机，它无须为货物和乘客提供空间，也不必提供进行长途飞行所需的燃料。

物理模型的进展大大超出了我儿时的实践。汽车制造商和空气动力学研究中心仍然利用风洞来研究不同车辆和飞机设计中所发生的湍流(turbulence)现象。船舶设计师在大的水箱和水池里拖曳或推进不同设计的船体，以便了解在各种波浪和风速的条件下如何提升航

速、增进燃料效率和提高稳定性。 物理模型在试验方面继续扮演着重要的角色，在下一章节我将要对它们作更多的讨论。 但是物理模型研究的成功与否，取决于一个试验在某个尺度中做得如何，在与真实环境不是完全一致的环境下，将其外推到真实气象条件下天空或大洋上的飞机或船舶。

数值模型

数值模型由描述系统功能和对系统行为进行定量预测的一系列数学公式组成。 当计算复杂到需要计算机来运算时，数值模型有时被称为计算机模型。 但是无论是用手工计算还是用老式加法机或手工计算器计算，或者用巨型计算机进行计算，所有数值模型的核心是定量联系组件和系统过程的公式。 开尔文勋爵通过求解热传导微分方程的方法计算地球年龄。 这一方法能够让他定量计算一个球体从最初热状态损耗热的速率，从而计算地球已经冷却了多长时间。

为了了解什么是数值模型，让我们从更简单和更直觉的事情开始吧。 我们以你的储蓄账户模型为例，每月你在其中存一笔款，账户内的钱会以给定的比率逐月变化地获得每月的利息。 有一些简单的数学规则让你能够计算出未来任何时间你的储蓄金额。 实际上，这些规则是一个关于你的储蓄账户的数值模型，一个代表了随时间变化的存款及组合的模型。 如果你想定期地提取现款来支付你的汽车保险或租金，模型会制作得更加复杂。

分期偿还几年欠下的家庭贷款的进程表同样是一个简单的数值模型的输出结果。 这一模型表明你借款的数额、你与借款者协商的利率、你希望用几年还清贷款；接着它会直接运用数学来告诉你每月必须支付的金额。 另外，它会告诉你每月支付的款项多少用来支付利

息，多少用来减少本金。在贷款的任何时候，如果你决定用从珍尼(Jenny)姑姑那里继承的财产或者用从办公室获得的一大笔奖金还清债务，那么这一模型将告诉你还剩下多少本金需要重新支付。

社会保障

美国人熟悉的另一模型是美国社会保障制度模型。这是一个经济模型，它同个人储蓄账户在许多方面具有共同点，但也有一些不同和更复杂的特点。一个不同点是，制度不仅能够预算在个人工作年间的存款，而且能够预算在以后退休年间的花费，使这种花费尽量与累积的存款水平相当。第二个不同点是没有付款的时间限制；只要人活着这种好处就会持续。对照起来，个人储蓄账户只有当平衡维持时才能支取金额。第三点不同，也是重要的不同，是社会保障制度拥有上百万的参与者，他们共同存款，同时制度从这一共同的存款中为年长者提供费用。我称此为模型的原因是社会保障制度的执行者、社会保障管理机构，在未来任何给定年份，必须对将有多少人参加工作和作贡献以及多少人退休和接受钱款做出估计。而且，他们必须对人们将来的寿命做出估计。此外，在未来的几十年内他们必须一直做这样的估计。因此，不难理解与社会保障制度的财政能力相联系的不确定性存在的原因。

当20世纪30年代引入社会保障制度时，为了确定每个工人必须为共同存款贡献多少，管理者对家庭规模、移民趋势、就业水平、退休年龄、预期寿命和未来利率做了有见地的猜测。那些估计准确吗？对于任何从专业角度仔细思考这些主题的人——人口统计学家、经济学家、保险精算师——都不会对“最初计算中用到的估计并没有与20世纪所展现的现实完全一致”这一现象感到吃惊。20世纪

30 年代的预言与随后 70 年的实际进展肯定是不同的。

在 20 世纪末的不多的几个 10 年内，对社会保障制度的供给开始落后于支取，其中的原因部分在于家庭变得更小，提供的劳动力更少；另一部分原因在于人们的寿命更长了。1935 年，在美国人口中，只有 650 万人(约占人口 5%)的年龄超过 65 岁。到 2000 年，超过 65 岁的人口已经增长到 3500 万，占到国家总人口的 13%。在 21 世纪初，已经到达 65 岁的美国人将有一半活着。就像对天气预测一样，人们试图对未来计划得越长远，预测就越变得越没把握和越不确定。

毫无疑问，在一个滚动的基础上，对于未来四五天的天气预报是精确的。能够取得这样的成功，原因在于无论在预测和现实之间出现多么小的差异，每天为未来做的计算都要进行调整以便减小这种差异。事实上，这种计算稍稍进行了修正以后，就会更好地对第二天的天气做预测。社会保障计算同样也经历了调整。当资金不足作为一种未来的可能性开始隐现时，受到国会批准的社会保障管理者，每隔几年要为工资增加一个基本的量以便为共同存款上税。近来，由于认识到 20 世纪发展的生命预期值在增加，因此能够从共同存款中支取全部救济金的人的平均寿命也慢慢地从 65 岁增加到 67 岁。甚至在今天，还有人在讨论是否应当增加共同存款的投资收入以便让共同存款和救济金进一步协调。

关于社会保障制度定量模型讨论的一个重要教训是，在对制度进行初始化时，人们无法预料使得预测模型对现实世界描述越来越不充分的所有不测事件。然而，有关未来的不确定性并没有阻止制度的引入，而且多年来，对制度的中期方向修正已经修改了制度并扩展了它的生存能力。在不确定性面前总是必须做出决定，由于有了社会保障制度的诋毁者归因于该制度的不充分性，对社会保障基本概念的

强有力的公众支持和对70年前在大的不确定性气候中建立的社会保障制度的赞赏就始终存在并继续下去。

模型需要修改，这样它们会更好地代表它们模仿的现实世界。模型的改进来自于实验，在实验中模型要面对不断变化的条件和刺激，其反应或者反映也要被仔细观察。在湍流的特定条件下，坠毁的模型飞机一定要送回制图板进行修改。在社会保障制度模型中，一个人要能运行“如果怎样”（what if）实验，以便了解需要什么额外的财政收入来表达在下一个20年平均寿命延长5年的影响。模型和实验之间的界限几乎是透明的，因此让我们武断地宣布结束关于模型的这章内容，接下去讨论实验吧。

注　释：

1. 我在第九章简要讨论了加州公用事业的取消管制，发生于预测出现错误的背景下。
2. 魏格纳的著作最早于1915年以德文出版（*Die Enstehung der Kontinente und Ozeane.*）。英译本于1966年由纽约的Dover Publications出版。
3. Reginald, V. Jones, *Most Secret War: British Scientific Intelligence 1939～1945*。Hamish Hamilton, London, 1978。亦由Wordsworth Editions于1998年以平装本重印。

第八章　如果……让我们来看看会发生什么

只要一个人还在提出问题，他就不会是一个真正的傻瓜。

——查尔斯·斯坦梅茨(Charles Steinmetz)*

“实验”这个术语就暗含不确定性，要不然如果实验的结果是确定的，那为什么还要做实验呢？实验的目标是了解一个系统的新东西、未知的东西或者所知甚少的东西。实验是模型的自然发展，因为无论是概念的、物理的或数值的模型，都是对系统的简化表达，对模型进行的实验有助于我们了解模型的优势和弱点。从结果来讲，模型中包含的简化在许多情况下可能不妨事，但是在特殊形式下，模型会非常脆弱。在这种背景下，即模型确认过程中，需要通过实验对它们进行检验。

就像模型的建立开始于一个概念——某东西如何被构建或起作用

* 斯坦梅茨(1865～1923)，1889年由德国迁居美国，后来成为美国通用电气公司的数学和电气工程师。——译者

的智力想象，实验也是开始于某种被称为“思想实验”的东西。这些实验是探索假说的结果或行动的可能路径的智力活动。爱因斯坦是一位坚定的思想实验的提倡者；他关于相对论的许多早期概念来自于他对宇宙所做的“如果他乘一束光旅行，宇宙将会是什么样子？”*的努力想象。

实验本质上是提出问题和寻找答案。一个好的实验会为一个适当的问题提供明确的答案。一个好奇的男孩也许会设计一个实验来回答“下落物体的加速度是多大”这样的问题。在他的实验设计中，他想象从一座高楼的顶层扔下一个保龄球。在扔之前，他将在下面每一层楼的窗户旁安排他的朋友，用秒表记录球经过每一个窗口的时间。当球下落时相邻两层所用的时间间隔是如何缩短的？对此问题的分析将揭示球在万有引力作用下获得速度的快慢程度。换言之，能够通过实验的再构造来检验一个假说或预测：牛顿爵士关于运动的第二定律预测自由落体的加速度等于下落距离的2倍除以所用时间的平方。如果观察显著偏离预测，那么实验者也许要对牛顿第二定律的真实性产生怀疑，或者一次又一次地进行实验来看看是否会得出同样的结果。后面的策略是科学家们检验彼此工作的长久方法。

其他的实验也许不是这样简单或直接。一位鱼类生物学家可能会问：“湖水中的杀虫剂会对小嘴鲈鱼的繁殖力产生什么影响？”然后她会在孵鱼卵的水箱里着手进行一项实验，其中每一个水箱杀虫剂的浓度是不同的。监视每一个水箱中鱼的后代数目和鱼的生存能力可能会提供一些相关的数据。但是在水箱中进行的实验会与附近由密集农业环绕的湖泊中发生的“自然实验”一样吗？除了杀虫剂流失之

* 可参看《爱因斯坦文集》，第一卷，商务印书馆，1976年，第1～42页。——译者

外，诸如酸雨和由上风向的发电厂产生的通过空气传播的汞等众多其他因素也会影响湖中鱼的繁殖力吗？ 湖中其他鱼类的存在和竞争会影响鲈鱼的成功繁殖吗？ 其成功繁殖又会取决于湖泊每年经历的温度范围吗？ 湖中鱼类的年龄分布会对它们的繁殖力产生什么样的影响？

生物学家也许回答了由她的孵化水箱实验提出的有限问题，但是却留下了自然环境中真实发生的更大问题没有回答。 自然环境正经历着一个自然实验，该实验条件受人类行为影响但不受人类行为控制，这些条件正在发生变化而系统也对此作出相应反应。 显然，在实验室里模拟自然界是很困难的。 与鱼类孵化实验结果相联系的不确定性比仅仅是在实验水箱中测量的不确定性要大。 这更大的不确定性来源于实验是否很好地表达了自然中正在发生的情况。

实验者设计实验的方式对于实验的最终成功或失败是至关紧要的。 常常设计一个实验来检验对于像“一个系统在各种条件下如何表现”这样问题的预见性想法。 更早我们称此为概念模型。 在鱼的孵化实验中，有一个假设：水中相对低浓度的杀虫剂对于鱼的繁殖力没有影响，唯一的实验变量是杀虫剂浓度。

关于系统行为的预见性想法，尽管在设计实验时几乎是必需的，但也可能成为实施实验和解释实验结果的阻碍。 医学和制药科学频繁地使用以双盲协议(double-blind protocol)著称的实验程序，设计该程序的目的是将实验进行过程中引入的偏见减少到最小。 在对一种新药的测试中，将要选择一群病人，他们之中有一半接受此药品，另一半接受安慰剂(placebo)[1]。 然而在双盲实验中，无论是病人还是实施测试的研究者，都不知道哪些病人接受了药品而哪些病人接受了安慰剂。 当然，通过先前对病人和治疗的仔细编码，将能了解到一些“谁接受什么”的信息，经过一段监护期以后，药品的效果能够从

所谓的安慰剂效应[2]中分离出来。这样一个过程实质上减少了因病人悲观地意识到他们被给予安慰剂或研究者鼓励已经接受实验性药品的病人所引发的对实验结果的偏见。

尽管预想几乎是不可避免的，但它很可能成为评估结果的绊脚石。如果实验揭示的结果同预想有相当大的差别那又会怎样？标准的科学传说是实验者被与先前假说相冲突的证据说服，然后改变他关于系统的概念。然而在现实世界中，时常会出现不情愿承认失败和轻易放弃个人假说的情况。也有一些实验者在暗示甚至证明系统运行方式的新概念的实验证据面前，执着地坚持他们最初的思想。[3]在科学上，有一个重要的警告："不要太依恋自己的假说。"

在地球内部

一些实验并没有提供一个直接的答案，它们只是缩小了可能答案的范围。在这些实验中，重要的结果是排除了对问题的一些提议性回答。例如，小孩子们经常问的、地质学家也努力回答的问题，"地球内部是由什么组成的？"得向凡尔纳(Jules Verne)*道歉，到地球中心的旅行实际上是不可能的。最深的矿井和最深的钻洞所达到的深度从现实来讲只是地球表面肤浅的小孔，我们从中得到的岩石样本代表的只是地球外层的物质。甚至通过火山喷发由自然提供的内部样本也只有最靠外层的大约数百英里。然而地球中心位于地表下几乎4000英里的地方。人类和自然对于极深地球内部的直接采样仍然是达不到的。

---

* 凡尔纳(1828～1905)，法国科幻小说家。主要著作有《从地球到月球及月球周围》、《格兰特船长的儿女》、《环游世界八十天》、《海底两万里》。——译者

然而并不是没有希望。我们能够从地震波（振动）了解到地球内部的一些东西。这些穿越地球内部的波，当它们又一次到达地表时，会告诉我们沿着这条路线遇到的物质的物理特性的一些情况。这些特性包括可压缩性、刚度和密度。于是地质学家面对的任务是确定什么物质在沿着通过地球内部的那条路线遇到的压力和温度的条件下具有那些特性。巨大的地球是由我们在地表找到的花岗岩、石灰石等岩石组成的吗?

用各种备选的岩石做实验，对其施加巨大的压力压缩它们，将它们加热到几千度，以此模拟地球内部的条件，帮助我们确定它们的特性是否和地震波通过地球作深层旅行时所遇到的物质显示的特性相匹配。经过这样的实验之后，作为候选的花岗岩和石灰石被轻易地排除了。无论使用多大的压力和多高的温度来对付这些岩石，它们都不会显示地球内部的深层特征。特别地，再怎么对它们压缩，它们也不会达到甚至只通过地球一半路程时的物质的密度。但是另一种叫作橄榄岩的岩石，尽管在地球表面不常见，却与地球外半部的特性匹配得很好。

然而，没有岩石能够满足那些从最深内部观察到的特征。对于这一区域，我们必须抛弃作为候选的岩石，去寻找完全不同的东西。诸如铁这样的重金属在岩石不满足测试标准的深度符合这些特征。因此，这些实验暗示的关于地球的大致图景是岩石组成的外半部（然而，原则上不是花岗岩或石灰石）包围着占据了内半部的金属核。

因此，对于地球的内部组成，我们是否真正了解？我们不了解。但是这些实验在解释地球不是由什么组成时是确定的。例如：除了地球表面一个非常薄的表层之外，组成外半部球体的岩石不可能是大量的花岗岩或石灰石，而且密实的内部根本不可能是岩石。通

过实验测试的材料之所以有资格成为候选岩石，是由于这些区域可能是由这些物质构成，但是并不能排除满足这些要求的其他物质。因而地球的组成是不确定的吗？是的，从地震和实验室数据（和许多额外的间接证据），我们能够就关于地球的主要组成作出自信而有见地的断言，但是在许多重要的细节上，我们很少能够确定。

关于岩石力学性质的实验室研究也被用于理解地震发生的物理基础。当地壳岩石断裂时，地震就会发生，此时允许破碎表面任一边的结构（地质学家称之为断层）滑落越过另一边。在实验室研究中，将小岩石样本放在巨大的液压设备中，对其进行挤压、拉长或扭曲直至它们破碎，然后观察需要多少必要的“折磨”能够让给定类型的岩石破碎。

能够将实验室结果迅速进行外推来预测现实世界的行为吗？在这方面有许多告诫。在现实世界中，通过自然过程对地球地壳的扭曲一般是发生在千百万年的时间尺度内，然而在实验室，即使一个耐心的实验者，这些挤压和扭曲却发生在几天或几个月之内。这些物质在自然中接受缓慢的“折磨”，与它们在实验室相对快地接受“虐待”相比，所表现的行为是不同的。在现实世界，与一场大地震相联系的断层在几英里之上也许才裂开，然而在实验室的实验中，典型的样本被拉裂时的尺寸只有几英寸长。在更大的自然环境中，岩石种类的异质性在确定断层可能在哪里发生以及何时发生方面很可能扮演着一个重要的角色，然而在实验室，可能更多的是样本的同质性。对于地质学家来说，这只是一种生活事实。他们在实验室对地质过程的努力探索已面对显而易见的事实，这就是地球并不很适合进入实验室，而且人类没有以自然进程的速度来从事实验的奢侈时间。通过小的简化实验室模型对大的自然系统的行为进行预测是一件充满着

不确定性的难事。

数值实验

数值实验的基础是对所研究的系统进行定量描述的数值模型。提供日常天气预报的数值模型运用流体力学的物理原理；国家经济的数值模型运用许多经济变量、政治变量和人口统计变量相互作用的经济概念。不必说，如果物理原理或经济概念是不完整的或不正确的，相应的对未来的预测也将是单薄的。从对复杂系统的理解是不全面的这个意义上来讲，未来仍然是不确定的。

因为数值模型存在于计算机内部，因此它们对“快速处理许多实验”负特别的责任，这些实验用于探索模型对于多种输入因素的反应。我们称这些实验为“如果怎样”实验。“如果我是20年而不是30年还清我的贷款，利息费用中的存款将是怎样”的数值模型，在一系列数值的“如果怎样”实验中成功地完成了每一步，阐释了许多成果，这些成果是许多不同输入计划说明书的结果。

在许多情况下，数值模型由于其成本低廉和多样性，在实验中已经代替了物理模型，因而它们能够在一个更广的变量范围内进行实验。人们在计算机内进行实验与在实验室中使用真实物质进行实验或者在风洞中使用飞机模型进行实验相比，能够进行更多的“如果怎样”实验。而且一些实验仅能在计算机上完成。我们不能通过直接的物理实验回答如下问题：“如果大西洋和太平洋不是被巴拿马地峡分割开来，地球气候将会如何？”但是我们能够创造一个大气和海洋循环的数值模型，这将使得在海洋和大陆的地理外形中随便选择进行实验成为可能。

美洲杯赛

每隔几年，全世界的帆船选手来到海边参加一个以美洲杯赛著称的重大比赛。这个比赛开始于1851年，当时只是在美国和英国的绅士水手间进行的精神竞争，今天已经发展成为一种包括公司赞助商和晦涩难解的技术规则在内的遍及全球的国际竞赛。不过，它仍然热衷于敏捷的航海技能以及在船体和帆的设计方面对技术优势永无止境的追求。正是在后一领域对于技术优势的追求之中，模型扮演着核心角色。

这一竞赛的早期历史中，新的船体设计实际上是在水中建造和进行测试以便了解在速度上获得了什么样的进展。但是这种反复实验的策略被证明是过度昂贵的（赛艇从来不是低廉的玩具）。后来，通过室内流体动力测试箱牵引或吹动的缩减尺寸的模型，代替了以实验为目的的全尺寸船体。今天，用于模型实验的竞技场在快速科学计算机里总是困难重重，在计算机里，流体动力学公式让虚拟的水与虚拟的船体发生作用成为可能。这样的研究是与在实验室或在自然环境下进行的实验形成对照的数值实验。计算机里的数值实验作为设计具有洞察力的源泉已经代替了水中的真实船体。通过数值模拟变化的风和海洋状况，测试船体和帆的外形来确定什么能够让一名熟练的选手弥补速度上的不足。最终，这样的实验结果必须接受海洋中自然条件下现实世界比赛的检验。

试验核武器

在依赖计算机模型和实验的方向上，更为大胆的步骤是应用于核武器设计和对它们的保存期限的评估。核武器的研制及其使用让第二次世界大战得以结束，自那以后，一个新的武器系列，既有更大

的，也有更小的，被设计和试验来确保它们完成预期的任务。包含成千上万枚不同设计的弹头的军械库已经通过原子能积聚起来，最初仅包括美国、苏联、法国和英国，但是今天，中国、印度、巴基斯坦，也许还有其他国家也已都在这一阵营中。到目前为止，除了在第二次世界大战中对日本使用了两枚原子弹，在战争中还没有使用过核武器。因为使用的缺乏，引起了一些对于“核武器在装配之后很长时间没有使用，现在是否能继续使用”的关注。因此，美国和其他可能拥有核武器的国家会引爆一个核装置来确定自从核武器制造以来在任何方面是否已经退化。在美国，这些试验在离拉斯韦加斯不远的内华达州试验场地下进行。对这些超出保存期限的武器的试验是现实世界的物理试验。

主要的核武器国家已经意识到核扩散的风险对国家安全的威胁也许比任何来源于新武器开发逐步增加的安全更大。因此，一个禁止核武器(新的和旧的)试验的国际条约，已经起草并置于国际谈判桌上。这一条约，就是著名的《全面禁止核试验条约》(CTBT)，于1996年由美国总统克林顿签字，但没有得到美国参议院的批准。发生于2000年的争取参议院批准的辩论中，主要的争端是，这些没有定期试验的武器在多年之后可靠性如何，是否所有组件继续工作。

条约的拥护者主张不再需要核武器的开发和试验了，因为对于武器开发的物理和工程方面建立计算机模型已经切实可行，这种模型将会排除对真实武器进行真实物理试验的需要。至于可靠性问题，一个热核武器的运行包括电子组件、机械组件和化学(非核)爆炸以及没有被条约禁止的试验。因此关于这些组件正在老化的效应能够继续被进行研究，替换品或者改进的版本能够不时地被安装到贮存的武器上。

尽管美国没有批准《全面禁止核试验条约》，但自1992年以

来，它已经遵守了条约的规定。那时美国在布什执政下，宣布暂停核武器开发和试验。在暂停前后，武器清单的可靠性已经通过美国能源部的储备管理计划进行评估。这一程序强调三个行为：(1)进行对贮存武器的监控来检测武器在非核方面的缺陷；(2)维修或替换组件以消除任何在监控时发现的缺陷；(3)开展对老化过程的基础研究，当关于长期行为的认识提高时能够重新设计组件。

关于老化的研究既包括实验室试验，也包括受缓慢的原子衰变影响的特性的计算机模型设计。与核电站所消耗的燃料棒的安全贮存和保留相联系的许多问题，也出现在核武器老化的研究之中。钚的半衰期为24 400年，它被使用在核武器中，同时也产生于核电站燃料的再处理过程中，钚在比实验室观察长得多的时期内，影响着外壳材料。这引起了对计算机建模和仿真的需求。

社会科学中的实验

如果像全球气候系统或热带森林生态系统这样的物理和生物系统在它们的复杂性方面使人畏缩不前的话，那么社会系统和人类行为的概念化和模型化方面的困难也同样具有挑战性。人们如何对人类贪婪、合作和利他主义的王国进行探索和量化呢？

社会科学家已经设计了简单然而非常有趣的数值实验，这些实验被作为计算机“游戏”执行。探索贪婪和公平性方面的游戏被叫做“分割赌注”。[4]这一游戏包括两名游戏者。有一笔钱，比方说100美元，放在桌子上。让一名游戏者决定他愿意与另一名游戏者分享这笔钱中的多大份额，然后将剩下的归他自己。然而，如果另一名游戏者拒绝这一分配方案，两名游戏者都将一无所获。

很明显，提供出价是有优势的，即使少于一半，也有着非常大的

诱惑力让另一名游戏者接受它。另一名游戏者会接受10美元，让第一名游戏者留下90美元吗？虽然人们会想象无论怎样出价都应该被接受，因为要不然两个人都会一无所获。但是研究显示在这一水平上达成协议非常不可能。一个35/65分割如何？也许可以。一个50/50的分割无疑会被第二名游戏者接受，但是还没有测试他可能接受更少的可能性。在这一游戏中，对于一名游戏者来说，什么是让他的奖金最大化的最好策略？当游戏在相同的两名游戏者之间多次进行时，该策略进展如何？当有许多对游戏者同时进行游戏时，该策略如何发展？

另一个计算机游戏，在理解人类如何进行合作和在什么条件下进行合作方面提供了很好的启示，此游戏被称为“囚徒困境”(Prisoners' Dilemma)。这一游戏是在监狱中进行的，在那里有两个偷盗嫌疑犯被拘留，每一个人被独立询问。任一个嫌疑犯不知道另一个人告知警方的事情。任一个被拘留者能够提供证据让另一个人牵连到盗窃案中。警察没有实证，需要通过一个或另一个嫌疑犯的帮助来使案件揭晓。

每一个嫌疑犯在不知道到另一个嫌疑犯说什么的情况下，他将怎样做？如果任何一方都没有告发另一方，他们都将获得释放并且分享偷来的钱。如果两者都告发对方，他俩将都进监狱。但是如果一个告发对方，然而另一个保持沉默，沉默的嫌疑犯将入狱而另一个将获得释放并且独自拥有所有的赃款。因此，为了让一方指控守口如瓶的另一方，就要用比两人都沉默时所得更大的报酬进行诱惑。但是如果另一个同样这么想，也控告他的同伙，那么两个人都受惩罚，没有一方受益。这一游戏只有4种可能的结果。

1. 两个囚犯都保持沉默，分享一半的钱；

2. 囚犯A告发囚犯B，但是囚犯B保持沉默。于是囚犯A被释放并且得到所有的钱，囚犯B进入监狱。

3. 囚犯B告发囚犯A，但是囚犯A保持沉默。于是囚犯B被释放并且得到所有钱，囚犯A进入监狱。

4. 两个囚犯都告发对方，都进入监狱。

如果我们称保持沉默为一个合作行为(C)，告发对方为一个背叛行为(D)，那么这4个选择能够被总结为C/C，D/C，C/D和D/D。这种困境显而易见：沉默会赢得自由和一些钱，但条件是你的犯罪的同伙也这么做(C/C)。指控对方是一种能够获得两倍金钱回报的赌博，但条件是你的同伙没有告发你(D/C，C/D)。如果两个人都告发对方(D/D)，那么两个人都被惩罚而没有任何金钱回报。你的选择是什么？你是合作还是贪婪？你对同伙的评估如何？如果你和同一个人重复面对这一困境，和这一游戏只进行一次的情况做对比，你的行为会有所不同吗？

在囚徒困境中，多个参与者彼此之间重复试验，从中会得出一些非常有趣的观察，那就是在同其他人相处时让受益最大化的策略。一个熟知的策略被称作“针锋相对”，其中第一个游戏者总是选择不告发另一个，除非在先前的游戏中第二个游戏者已经告发了第一个游戏者。在那种形势下，第一个游戏者通过下一次背叛获得平衡，有效地使用了“针锋相对”。如果第二个游戏者再次从失利中后退，那么第一个游戏者也回到了合作的姿态。这一策略具有一般合作策略的特征，但是如果遇到一个非合作行为时，就会直接报复。在单一的失利和报复事件之后立即回到合作姿态等同于不记仇。这一游戏在因特网上长期的试验中已经被千万人玩过，[5]而且结果被广泛分析，用于知晓在人类相处过程中合作和贪婪是如何发展和进化的。

“针锋相对”就如它的字面意思一样简单，也已经被证明是最大化长期收益的最有效策略。

“囚徒困境”的变体已经被发展用于诊断婚姻内外的关系，探索类似“如果一些人认为他们的事不会被发现的话，那么他们会不会欺骗他们的配偶？”这样的问题。社会科学家现在用数值模型和计算机游戏来表达类似的下列话题：人们如何对动机和团队利益进行反应？人们为什么会拖延？什么情况下犯罪率会下降？[6]这些模型中的一些，比如说“分割赌注”和“囚徒困境”，是非常简单的，然而，其他模型却包含着大量的统计学和数学。它们的共同点是都集中于人类行为的复杂性和精细性。

无意中做的实验

不是所有的实验都是在受控的条件下精心设计和执行的。一些“恰恰发生”在不经意的情况下。然而从不经意的“实验”中可以学到许多东西。1990 年 5 月，一艘装载亚洲货物去北美的集装箱轮船在北太平洋遭遇到非常恶劣的海况，一些装着运动鞋的钢铁集装箱从甲板上掉到水里并且裂开了，有约 80 000 只鞋子被扔进了大海。[7]这些鞋子小心翼翼地漂流着，一路上受到洋流的推动，直到它们沿着北美西海岸登陆。就这些漂流的鞋子来讲，知晓鞋子进入海水中的地点和它们最终到达的目的地，能够让海洋学家对洋流进行描绘和更好地理解洋流的宽窄以及由它们创造的旋涡是如何导致鞋子等示踪物的传送和分散。实际上进入海水的鞋子沿着加拿大和美国的海岸线到达了 500 多英里外的地方。这样一个实验也许已经被科学家想到并实施，但事实上整个事件是不经意的。实际上，这一实验在无意中已经又被重复了两次。一次是在 1992 年 1 月，装有 29 000 个浴缸

玩具的集装箱在北太平洋被冲出船外，放出蓝色海龟、黄色鸭子、绿色青蛙到阿拉斯加锡特卡附近的海岸上。第二次是在 1994 年 12 月，34 000 双曲棍球手套“在海中失落”。

最近我去医院访问了一个朋友。在进入她的房间之前，我被要求穿上一件消过毒的长袍，戴上乳胶手套。我被告知在医院里散布着耐抗生素的细菌，我的朋友由于身患癌症已经很虚弱，在这些抗药性甚高的微生物面前就更加脆弱。这些微生物已经通过另一个不经意的实验获得了对抗生素的抗性——原本很容易通过其他方式治疗的疾病，却被漫不经心地滥开抗生素处方。当不必要使用抗生素却使用了抗生素时，医生们已经在不经意间让这些有害的微生物瞥见了医药库。细菌会利用这一机会通过进化机制来对付抗生素，因而在微生物与医药之间就会发生不断的战争。在兽医学中已经观察到类似的效应，在那里原先已被证明在对抗特定疾病时很有效的抗生素，牲畜已经不再对其产生反应了。类似地，为准备去热带地区的旅行者和士兵开的抗疟疾的预防药物已经逐渐地丧失了效力，原因在于疟原虫已经产生了对这种新武器的抗性。

臭氧洞

另一个大规模的不经意的实验开始于 1929 年，那时，为了取代当时令人不太满意的致冷剂——氨(NH_3)和二氧化硫(SO_2)，杜邦公司开始将一种新型的无毒家用惰性致冷剂推向市场。前者虽是过得去的致冷剂，但是不仅易燃而且(或者)有毒性，在家庭中使用会引起危险。后者，这一新的杜邦致冷剂属于氯氟碳(chlorofluorocarbons，简写为 CFCs)化学品家族，是氯、氟和碳的化合物。这些化合物，除了具有极好的致冷特性，还有许多其他令人满意的特性：如果泄漏

出冰箱，不会燃烧或爆炸，对健康没有危害，不溶于水。

第二次世界大战后，发现了CFCs的其他用途。这种作为气溶胶能够被有效配送的产品，被证明是很好的推进燃料。毕竟它们与和它们混合的物质不发生反应，也不溶于这些物质。在压力下它们能够被压缩，当释放时能够用作推进燃料。很快，它们成为头发喷射剂、臭虫喷射剂、空气清新剂、喷绘涂料和作为气溶胶喷到目标的许多其他产品的推进物的选择。不久，CFCs也用于诸如聚苯乙烯泡沫塑料和聚苯乙烯之类的刚性泡沫产品的包装和绝缘。在家用冰箱之外尤其是在建筑和汽车空调领域，其致冷特性发现了新的用途。CFCs的确是一个了不起的产品，服务于非常广泛的各种需求。然而，当它们广为普及时，一种不经意的副作用逐渐显现出来：CFCs分子已经上升到地球的平流层，在那里它们成为破坏平流层臭氧的主要元凶。

臭氧是氧气的一种特殊形式，其中，在单个分子里有三个氧原子。普通的氧气占了地球大气组成大约20%，仅包含两个氧原子。在平流层臭氧是通过自然过程产生和被破坏的，经过很长时间获得一个非常稳定的平衡浓度。虽然其浓度在大气中甚至远远不到1%，但它扮演着非常重要的环境角色：它过滤掉许多来自太阳的紫外线。这种射线对健康有害，因此平流层中的臭氧，即使浓度很低，仍然是一种非常有益的成分。

不幸的是，正是那些使CFCs成为有吸引力工业产品的特性，也对地球平流层的臭氧造成始料未及的破坏。每按一下喷壶按钮，就会将CFCs放入大气。每一个废弃的电冰箱、每一辆装有空调的废弃汽车最后都会泄漏CFCs进入大气。惰性的CFCs一旦进入大气，是不能同其他大气成分化合的，也不能被雨水冲刷，因为它们不溶于

水。CFCs累积在大气中，受风和天气作用，最终会有一些到达臭氧存在的平流层。一旦进入平流层，它们会打破存在于臭氧产生和被破坏的自然过程之间的平衡。

在平流层，来自太阳的紫外线会提供足够的能量，从CFC分子中分离出氯。这些对氧充满贪婪喜好的游离氯，发现臭氧分子并进行攻击。它们会拖掉并捕获三个氧原子中的一个氧原子，因此臭氧分子降级变成一个普通的氧分子。然而，拥有新捕获来的氧原子的氯原子是短命的。单个漫游的氧原子群，发现被氯原子包围的氧原子，会拽出氧原子并与它形成另一个普通的氧分子。这时刚刚被剥夺掉氧的氯原子，会寻找并找到另一个臭氧分子进行拆分。因此单个氯原子会变成一个重复的冒犯者，在一段时间内捕获许多臭氧分子。由于这一过程没有得到控制，因此这会很快导致灾难性的臭氧损耗。然而，氯也能够稳定地与其他分子结合，从而离开与臭氧的战争。不幸的是，这一隔离氯原子的过程进行得特别慢；单个的氯原子在它最后被捕获或从平流层被清除之前，会回过来成千上万次地攻击臭氧。

由氯原子造成的臭氧破坏在南极洲的南极点区域尤为明显。在那里年度基础上的平均温度是寒冷的－45 ℃（－50 ℉），在阳光彻底消失达6个月之久的极点冬季就更加寒冷。这种极度寒冷促进了平流层中微冰晶体的形成。经过漫长的冬季黑暗之后，当阳光返回时，这些晶体的表面为反应的进行提供了场所。这些反应是：当返回的阳光提供能量时，将带有惰性氯的化合物转变成其他能严重破坏臭氧的物质。这些物质很快会对平流层的臭氧造成破坏。可以说，这是南极洲上空发生的一个惯例。

自从20世纪80年代早期，暴露于紫外线的南半球更大区域，南

极洲上空的臭氧破坏逐年增加。自1987年以来，根据《蒙特利尔议定书》这个国际条约，CFCs产品实际上已经被淘汰了，工业化学家已经采取措施来寻找扮演许多重要角色的CFCs的替代品。然而，虽然进入大气层的CFCs在逐步减少，但引入CFCs所带来的未曾预料的效应将持续到本世纪下半叶。

认识到"CFCs正在引起南极洲平流层臭氧的破坏"这一现象是一个科学探索故事的结果，这一故事最终导致发现这个相关过程的科学家获得诺贝尔奖。但就是这一故事，在前进的道路上遇到许多怀疑者，最初也引发了许多抵制。部分抵制的出现是因为在原因和结果间存在着很大的概念上的差距。有人认为普通的人类活动，例如像头发喷雾剂的使用，怎会对地球大气产生显著效应，这无异是一个异想天开的概念。带着一个大的星球的微小居住者的人类形象，我们已经成长，这个形象显示自然力比我们能够做的任何事情强大得多。我们已经进化到对于我们能够看到、听到、摸到、嗅到的事物和紧靠我们的环境中也许是一个威胁的事物产生警觉。在20世纪，意识到人类活动的缓慢积累，能够作为总体以同样对我们形成威胁的重要方式影响整个地球，这是智力的延伸。

关于CFCs和臭氧损耗的不经意实验具有那种让人惊醒的有益效果，它使科学家和许多其他人意识到人类已经对全球环境造成的总影响。发现损耗并确认原因的科学研究以及最终导致朝着补救采取国际措施的政治反应，在某种程度上已经为"世界所面对的一个更大的不经意实验——全球气候变化"做着准备。

变化多端的气候

也许，人类历史上最大的不经意实验——全球气候变化，当前正

在进行中。在这一实验中，兴趣的焦点在全球气候系统如何应对大气中日益增长的二氧化碳水平。在大气组成表中，二氧化碳排在第四的位置，远在氮和氧之后，氮和氧两者加在一起占了大气总体积的99%。二氧化碳甚至排在剩下1%中约占十分之九的氩后面。在1750年，恰逢工业革命进行当中，二氧化碳占据了大气体积中每百万单位中大约280个单位(ppm)，或者大约0.028%。今天，在经历了两个半世纪的燃烧化石燃料以便为全球工业经济供能之后，大气中的二氧化碳水平已经增加到约380 ppm。而且，实质上在有关未来全球人口、能源使用、技术变革的每一份合理方案内，二氧化碳浓度在21世纪下半叶早期将要增长到500～600 ppm左右，浓度大约为工业革命前水平的两倍，这是一个比过去50万年以来地质记录中所见的都大的水平。

我们需要关注占大气总体积的一小部分的这一微小大气组成的变化吗？回答显然是肯定的。无论体积大小，二氧化碳通过作为温室气体的行为，在调节地球表面的温度和气候中发挥着极为重要的作用。温室这一术语指某种大气气体捕获地球表面发射能量的能力。二氧化碳、甲烷、水蒸气、[8]CFCs和以痕量存在于大气中的其他气体成分显示了这种捕获特性。而且我们应该为出现这些温室气体而感到高兴，因为如果没有它们，地球表面温度将远在水结冰的温度以下，我们的星球家园将成为一个冰库。在地球上我们之所以拥有海洋、湖泊和河流，而不是冰河、冰冠和海冰的覆盖，就是因为温室气体的存在。此外，我们知道地球在它的早期就已经有这些温室气体，因为在整个保存于水中的沉积岩的地质记录中，有充分的证据可以证明。水而不是冰的出现，贯穿大部分地球历史，需要持续时间很长的温室地毯。[9]

最近几个世纪，通过燃烧煤、石油和天然气，大气中增加的二氧化碳正改变着地球的自然温室，使它作为一条覆盖星球的毛毯更为有效。就变化的海平面、减少的冰层、变换的降雨模式、极端天气事件、植被和农业的变化来说，地球的变暖在即将到来的时期将如何作用，还不完全清楚。但是许多变化已经在进行中，即便是补救措施马上开始，进一步的变化也是不可避免的。地球气候是慢慢的、递增的改变还是突然的、灾难性的改变，对这一问题的回答充满着错综复杂的不确定性。

公众对有关这一“自然”实验的持久而激烈的争论的反应，最初只是一种怀疑，这种怀疑与人们最初对于臭氧损耗的怀疑基于相同的原因。许多人发现以下现象是难以想象的：每个人的日常活动——为房间供暖、为房间提供照明和驾车——能够在大规模且强大的自然过程中扮演重要角色。此外，那些备受影响的工业开展了一些精心策划的、自我服务的公共关系运动，让公众认为科学是不成熟的而且也是缺乏说服力的。

已经被科学家接受但是同样受到人们怀疑的观点是，二氧化碳在地球上已经逐年增长到长久以来未曾见过的水平。不再会有有关“地球是否在变暖？”或“人类制造的二氧化碳是否正发挥着实质性作用？”等问题的许多科学争论了。现在讨论集中于“这一实验的结果会是什么？”、“我们期待什么样的结果？”、“要对抗或者适应变化的气候，我们该做些什么？”这些问题。我将在本书最后一章较为详细地回到这一伟大实验。

在地球出现人类之前的很长时期内，自然已经运行了许多这样的实验。没有一个我们能够重复，但是我们能够观察它们的结果。大陆已经从热带纬度漂移到极地区域，以前的海洋盆地已经提升变为绵

延山脉，变化的海洋环流变更着地球的气候。地球上的生命，伴随着由地质过程引起的环境变化导致的物种出现和灭绝，已经具有许多转折点。我们只能够通过研究这些自然实验的结果尽力理解这些实验；我们无法重复这些实验，也没有能力去改变实验的条件。

在历史和地质意义上重建过去，包括提出问题和寻求答案，是多种实验的一种。但它不是我们设计或控制的实验；它是在我们尚未出生时就已经进行的一个实验，我们只看到了它的结果。我们想要理解产生目前我们所见状态的过程和环境，但是我们将遇到有关“我们重建的任何事件是否是实际上发生的事件”的不确定性。在下一章我们将深入钻研一些有关“重建过去”的特殊类型的不确定性，并了解是如何适应这些不确定性的。

注　释：

1. 安慰剂是一种无活性成分但能影响病人健康状态的物质。
2. 对于安慰剂是否真的有效果最近有人提出质疑。参见 the *New York Times*，24 May 2001。
3. Harry Collins and Trevor Pinch 对预想影响实验解释的方式在下文中做了较为广泛的讨论，*The Golem: What You Should Know About Science*，2nd edition，Cambridge University Press，1998，192 pp。
4. 分割赌注也被称之为拿得起放得下或极端主义游戏。
5. 在因特网搜索引擎中输入“囚徒困境”将为你提供许多机会了解这一正在进行的试验。
6. *The Wall Street Journal*，27 April 2001；*the New York Times*，27 November 2001.
7. Ebbesmeyer，C. E. and Ingraham，W. J.，Jr.，Shoe Spill in the North Pacific. *EOS Transactions of the American Geophysical Union*，vol. 73，n. 34，pp. 361，365，1992.
8. 大气中水蒸气的浓度是随地区而高度变化的，其变化取决于大气的温度。
9. 在地球早期历史中水蒸气对地球表面温度的形成也许发挥了重要的作用。

第九章　重 建 过 去

你能向后看得越远，你就有可能向前看得越远。

——温斯顿·丘吉尔(Winston Churchill)

像考古学、地质学和天文学这样的历史科学，其不确定性表现为一种被称为“非唯一性”(non-uniqueness)的特殊形式。当试图理解事情为什么以它所运行的方式发生时，科学家们必须努力重建事件发生的环境，并对支配事件的过程做出假设。但是当我们努力重建一个历史事件时，我们面对的是一个不完全的记录。运用这种对过程只有部分理解的不完全记录对事件进行解释，也许有不止一种解释方法。换句话说，我们必须生活在非唯一性的阴影之下。在任何给定的时间，不完全的证据也许承认许多种解释，但到后来某个时间，额外的证据可能会排除一些可能性。

处理过去事情的不确定性是地质学家的一种生活方式，他们在重建自然历史的工作中总是用半块板或者不足半块板来工作。自然界并

不是一个细心的管理员，随着时间的推移，结果必然是很长时间以前发生的事情的记录总会变得不全面和不完整。在他们理解和解释不完全信息的努力中，地质学家总是运用少数与解释他们的观察相关的临时假设来工作。这种思考的方式被钱伯林(Thomas C. Chamberlin)奉为神圣，他是杰出的地质学家、19世纪末20世纪初威斯康星大学的校长，后来成为美国科学促进会的主席。在1890年，钱伯林发表了题为《多种工作假说的方法》[1]的文章。在文中他阐述了“不把一个人用于解释的全部鸡蛋放在一个篮子里”的观点。相反地，他主张，当一个人跑到一个黑暗的小巷时，应退回来并且选择另一条追求之道，越快，效果就越好。他称手头有很多替代物能够促进批判性思考，防止科学家形成思维惯性。钱伯林的观念在一位专家的名言中得到共鸣：“再没有什么比我们只有一个念头的想法更危险了。”

飞机掉下来……

近些年一些空难已经提醒公众有关飞行的危险性，虽然要比在汽车事故中死亡的风险要小得多，但也不是无限小。1996年环球航空公司(TWA)800航班的坠毁，2000年法国航空公司协和式飞机的坠毁，还有2001年美国航空公司587航班的坠毁——这一切已经经历了广泛的调查，而且在重建和理解导致它们各自坠毁的事件中，继续存在着不同程度的不确定性。

由于TWA800航班事件发生得早些，对它的调查时间要比对其他坠毁事件的调查时间要长些，这一事件为讨论重建过去事件的不确定性提供了一个很好的例子。1996年7月17日，飞机在离开纽约肯尼迪国际机场前往巴黎的途中起飞不久，在长岛上空爆炸坠入大西洋，机上所有人遇难。对此，人们提出各种假说来解释这一事件，证

据虽然不完全，但仍然从海上和其他地方收集以评估每一种假设的可能性。整个事件笼罩在关于究竟发生了什么的巨大不确定性之中。

随着1988年泛美航空公司103航班在洛克比上空爆炸，苏格兰的燃烧至今仍深烙在公众的回忆里。一点也不必惊奇，早期的解释包括关于恐怖主义活动的各种假说：放到飞机行李架上的一颗炸弹；放在躲过安检屏幕侦查的乘客舱的一颗炸弹；与安全员同谋私运到飞机上的一颗炸弹；放在餐用服务容器中被带上飞机的一颗炸弹；被一名清洁工藏在飞机上的一颗炸弹；一次使用地对空导弹的恐怖袭击。

第二类广泛的假说集中在飞机的机械故障方面：飞机的爆炸性减压；飞机上厨房内部的火苗；发动机起火导致燃料箱爆炸；因静电或短路引起的燃料箱的爆炸。而第三类可能性包括一些不大可能但也有微小可能性的原因：被闪电击中；与小飞机的碰撞；与一群在高空飞行的天鹅相撞；或者导致飞机解体的剧烈空气湍流。

调查通过许多渠道进行。飞机几乎是通过从洋底找回的残骸完全重建的。对命运不祥的飞机的雷达追踪已被仔细检查。地面上的目击者和附近其他飞机的飞行员提供了证词。安检员、行李管理员、餐饮服务员、清洁工都被讯问。对相似的燃料箱进行实验，以确定爆炸在什么条件下将会发生。4年后，经过其33年的历史上最长和最昂贵的坠毁调查，美国国家交通安全委员会发布了一份关于TWA 800航班坠毁可能原因的报告草案。[2]对各种解释经过深入地分析和评估后，国家交通安全委员会否定了各种各样的恐怖分子假说。类似地，它逐一驳回了不可能的事件。

潜在的机械故障成为调查的焦点。最终强有力的证据指向一个机翼中燃料箱的爆炸，极具讽刺意味的是，此燃料箱除了装有燃料蒸气和大气中的氧气组成的易爆混合物之外，几乎是空的。这种混合

物非常不稳定而且易燃，因为在起飞之前，飞机已经在烈日炎炎的停机坪上暴晒了大约3个小时，提高了燃料箱内部的温度。然而，究竟是什么引起空燃料箱的爆炸，仍然没有明确的、得到一致认可的答案。燃料规管线路的绝缘处受磨损是主要的疑点。

不确定性存在于我们对TWA800航班悲剧的理解中，而且将一直存在。大约有10%的飞机残骸留在海洋底部，永久消失了。我们无法采访飞机工作人员来了解爆炸后命运攸关的几秒钟他们经历着什么。我们也不能让飞机在受控的实验条件下重新飞行来验证关于爆炸原因的一些想法。然而，即使有着一系列被承认的不完全的证据，可能的解释范围已经被大大缩小了。现在一些解释与国家交通安全委员会在深入调查中拒绝的许多其他解释相比，有着更大的成为真实的可能性。然而，阴谋论仍在因特网上大行其是，包括宣称政府掩盖国际恐怖主义。这些理论的支持者在许多方面是“真实的相信者”的榜样，无论何种证据他们都有自己的想法。没有什么事实或逻辑推理会让他们改变信仰。因为他们没有能力或不愿意接受让他们最喜爱的信仰站不住脚的证据，因此我们必须把他们引入讨论中的不确定性放在一边。

不确定性也不是没有益处。由于我们不知道燃料箱爆炸的确切原因，我们被迫接受一些可能性，采取补救措施来表达这一多样性。是燃料箱太热了吗？也许飞机制造商能够安装更好的绝缘设备来保护燃料箱不受停机坪的热浪的影响。是燃油蒸气和大气的混合物太不稳定了吗？也许在燃料消耗时往油箱内注入惰性的氮可减少这种不稳定性。是燃料规管线路的绝缘层太薄了吗？一个较厚的隔热罩也许会有帮助。对于协和式飞机坠毁的调查显示飞机燃料箱非常脆弱，会被爆裂轮胎的碎片刺破。美国航空公司587号航班的尾翼故

障导致人们将注意力集中在碳纤维合成物的耐久性上。对每个悲剧原因的调查从一开始就隐藏着不确定性。但在每一个案例中，调查揭示了一个先前未曾预料到的设计缺陷，并已经导致设计和材料的改进。简而言之，最初的不确定性激发了创造性思维；最终，商业飞机会比以前更加安全。

TWA 800 航班的故事在哪些方面有助于我们理解科学的不确定性，尤其在重建过去方面的不确定性？历史学家、地质学家和考古学家例行公事地表述发生在很久以前的事情，他们总是必须处理一个不完全的证据。在理解人类的历史或者地球的起源和随后的进化方面，虽然说不上取得什么进展。然而，随着额外证据的累积和对概念进行的修正，我们已经了解了许多并且将会继续了解更多。即使一些假说被扔在一边，然而关于实际上发生了什么的不确定性仍将存在。

在法庭上

不完全证据、不准确证据、互相冲突证据成为准则的另一个发生地是法庭。在审判大厅，重建过去是经常发生的事情。谁做的？这一问题及其无数的变量在审问、回忆、推测、法庭辩论或者实物证据、心理状态和更多的方面被表述。每一诉讼的案例都涉及关于发生了什么的具有竞争性的说法（在科学上，就像本章前面提到的，这被叫做多种工作假说）。

不完全证据？就像我们在著名的麦克维（Timothy McVeigh）* 案件中了解到的，麦克维是 1995 年俄克拉荷马城政府大楼爆炸的犯罪

* 是指 1995 年美国俄克拉荷马城炸弹爆炸事件，该案件造成 168 人死亡。该爆炸案的元凶为麦克维。——译者

者，联邦调查局在审判之前未向辩护律师出示数千页的证据文档。在更通常的环境中，一个遗嘱的附件也许已经不见了，展示给法庭的婚前协约仅包括她的资产，或者杀人犯的凶器和逃跑的轿车再也找不到了。

不准确证据？一个关于来自杀人嫌犯武器的子弹的弹道试验无意间被贴上来自不同手枪的标签。一个证人证明逃跑车辆的牌照是COR-134，但事实上是C0P-1B9。而且随着犯罪和审判之间时间的流逝，记忆变得模糊，证人可能死了，证据也消退了。

冲突的证据？原告证人证实，在犯罪时间内被告在犯罪现场，然而被告证人证实被告在2000英里外的地方和他们在一起。犯罪现场的指纹是被告的，但是杀人凶器上的血迹既不是受害者的也不是被告的。一个公司的总裁被卷入一桩产品诉讼案中，根据一些内部档案，他是知道这个公司产品中存在安全危险的；而根据其他档案，他并不知情。

法官和陪审团必须经常对竞争性的说明、不完全证据和错误证据、完全冲突和经常虚假的证词进行详审。他们仔细倾听陈述，权衡和评估证据，并做出决定。陪审员不允许随意寻找更多证据或者参加冗长的研究，这种选择可能会澄清重要而含糊的论点。他们不能为了等待可能会减少围绕案例的不确定性而推迟做出决定。他们要做出的是一个“合理的怀疑之外”而非百分之百确定性的裁决。

陪审团曾经犯过错误吗？我们知道他们犯过。但是我们也知道判罪有时和新的物证是背道而驰的，它放弃证词或者放弃随后其他人的招供。我们接受法制系统的不完美，原因在于宪法赋予了与我们地位相等的陪审团加速判决的权利带来的更大好处。我们承认法庭是一个不整齐的场所，在那里所做的决定通常是正确的，但是偶尔也

会犯错，那些决定是在不完全证据、不准确证据或互相冲突的证据面前做出的。

在水下

对能够处理不完全、不准确而且有时是冲突的信息所带来的好处的一个奇怪但又很有趣的确认，是从第二次世界大战后对成功的潜艇艇长所具有的品质研究中显现出来的。在这项研究中，确定一名成功的艇长很简单：在战争中活下来的艇长是成功的，而在战争中死去的则是不成功的。研究在许多因素之间进行了调查，包括什么样的教育背景会引起他们成功或不成功。这项研究得出的结果是：在诸如地质和经济之类的领域受过培训的人更容易存活，然而数学家或理论物理学家在战争中更倾向于死亡。

为这一结果提供的解释是：在水下攻击，潜艇人员工作在以不完全信息、不准确信息和冲突的信息为特征的不利环境中。声学的听力装置提供了有关敌人位置、敌人活动的一些信息，而且当靠得不是太近时，爆炸的深水炸弹能够提供一些敌人攻击计划的信息。地质学家、经济学家本身是习惯于用不完全数据进行工作的科学家们，更能够合理而准确地阐明正在发生什么和采取什么样的规避行动。相比较而言，数学家和物理学家，习惯了根据定义好的公理和规律进行高度结构化的科学思考，反而不能更好地处理这些观察。这些观察在严格尺度内不容易匹配，或者更糟的，看起来还相互冲突。潜艇战是一个非常混乱的形势，经常掉到数学家和物理学家习惯的结构化逻辑的外面。当他们对形势的不完美、逻辑结构的缺点或漏洞烦恼时，他们被一枚深水炸弹所击倒。

数学家和物理学家的致命缺点是他们本能地厌恶那些与他们熟悉

的理论框架匹配得不是很好的问题，而且也厌恶那些留下太多的选择来评估的不完整信息。他们习惯于解决具有一个正确答案而不是具有许多可能答案的问题。诸如地质学和经济学之类不太成熟的科学，就很少有这样的限制。因此，对于具有这样背景的艇长来说，他们更多的是靠直觉而较少靠分析。因此随着事情的发展，他们更容易幸存下来。

2＋2＝4，但是“?＝4”意味着什么?

数值模型通过缩小通向历史性结果的可能路径的范围帮助我们理解过去。科学家在运用计算机研究事情是如何进展以及为什么那样进展时，有时运用一种被称为“反演”(inversion)的数学运算。“反演”模型让我们在一种定量的基础上进行重建。无须对“反演”的数学基础进行深入钻研，人们能够轻易地把握这一方法的逻辑。考虑熟悉的问题：2＋2＝?。在该问题中我们假定两个数(2和2)和一个来组合它们的规则(＋)；通过使用这些符号的常用含义，绝大多数人会得出容易接受的答案“4”。当这两个数和我们操纵它们的规则是直接而清晰的时，结果就会是唯一的。我们称这种计算类型为“向前”(forward)模型，在此模型中，我们提供了配料和配方，获得一个唯一的结果。

现在让我们考虑这一问题的一种变体，仅规定“?＝4”。这是一个*反演*问题，一个重构问题，这里我们给出了一个结果4，要求确定是什么导致这一结果。立刻会有人抗议，“这是不公平的，有许多方法来产生4的结果！没有‘正确的答案’。”的确，这里有许多方法产生4：5－1；3＋1；8÷2；2×2；16的平方根，当然还有我们最喜欢的2＋2。而且我们不必把我们的选项仅仅限制在增加整

数。3.47 + 2.85 − 2.32 如何？ 你也可以得到同样的结果。 数学家告诉我们实际存在着无限方法来产生这一给定的结果，有无限多的说明可以解释这一给定结果的存在。 因此，我们做些什么？ 放弃我们无用的手吗？ 抱怨问题不可解决就放弃吗？ 不，我们要尽力运用能够帮助我们缩小可能性范围的其他信息。

根据一个现实世界的例子，想象一个盒子放在桌子上，我们被告知盒子里装有 4 美元。 要求我们确定盒子中的纸币和硬币是怎样进行组合的。 这是关于“? = 4”问题的一个例子。 我们如何缩小可能性的数字范围？ 首先快速晃动盒子，显示里边没有硬币哗啦哗啦响，这样我们只需要处理纸币。 对此问题的仔细思考能够让我们得出结论：我们正在处理加法的过程，也就是说，我们的问题是纸币如何组合加起来才为 4 美元？ 下一步，我们回想面值不足 4 美元的钞票只有 1 美元和 2 美元。 现在我们能够取得一些进展。 仅有三种可能：两张 2 美元的纸币、四张 1 美元的纸币、或一张 2 美元纸币和两张 1 美元的纸币。 对于“? = 4”的这样一般问题的无限多种回答，通过将加法知识运用到手头的特定问题的方式，已经减少到只有三种回答。 一个人通过运用从美国财政部获得的信息能够进一步推断货币流通中 1 美元纸币和 2 美元纸币之间的相对充裕性，并对三种解决方案中不确定性做概率估计。 尽管不确定性依然存在，但是它已经大大地缩小了。

找到一个问题可接受的解决办法的另一个方法是穷举法（brute force method），此方法主要是试验多种临时的答案并检查每一种答案是否有资格成为可能的答案。 例如让我们研究“? = 25”的问题。 此例子，就像我们先前的例子，有无限多的解决办法。 但是如果我们在这一过程上加一些额外的限制，我们就能够缩小可能的解决

办法的范围。让我们限定求的是两个正整数的和（整数，不是分数），接着给计算机编程序来试验许多不同的组合，并将加起来和等于25的那些数字一对一记录在列表中。这样做，实际上是让计算机构建一大堆“向前模型”，并且对每一个模型进行检验，看其是否适合作为解决问题的一种办法。

计算机很少抱怨这种单调的工作，但是我们通过以下的方法能对计算机有一点点帮助：很明显，每个数字必须小于25，否则它们的和不可能是25。因此，我们告诉计算机通过取1和24之间的数来对这些数字进行限制。而且，我们知道25是一个奇数，因此两个偶数，或者两个奇数加起来不可能为25。因此当计算机随机地生成两个数进行相加时，我们告诉它除非这两个数是一奇一偶，否则没必要把这两个数加起来。然后我们让计算机不受约束，运用指令重复使用这种操作5000次。

每当计算机找到一对这样的数字时，它就把它们加到可能解决办法的列表中。但是由于每次试验时并不了解前面的试验，因此完全有可能在5000次试验中不止一次地找到相同一对数字。因此在把这对数字加到列表上之前，我们让计算机检查这对数字是否已经存在于列表上。如果是这样，就记录下这个数字重复发生的次数，然后接着试验另一对数字。在5000次试验后，计算机打印出如下的一个列表：

1＋24	（4次）	7＋18	（5次）
2＋23	（7次）	8＋17	（4次）
3＋22	（3次）	9＋16	（3次）
4＋21	（6次）	10＋15	（6次）
5＋20	（5次）	11＋14	（4次）
6＋19	（4次）	12＋13	（2次）

在5000次试验中，计算机发现加起来和为25的整数组合共有12对。此外，它还发现它们之中每一对重复的次数，并相当自信不可能错过任何的组合。尽管当随机选择数字时，肯定有可能错过一种或多种对于问题的解决办法。如果我们只取100次数字，将很可能错过一些满意的组合。试验的次数越多，你将发现更多的甚至全部可能答案的概率就越高。但是在一个实际问题中，你将不能确切地知道你没有忽略一些解决办法。我们运用额外的信息大大缩小了研究的范围，有助于接近我们有理由认为解决办法可能会出现的数字的范围和类型。

在现实世界的更广泛的背景下，选择到哪里去寻找解决办法是寻找答案的重要组成部分。就像探测金矿一样，如果你知道一些关于先前发现的金矿的地质环境，那么至少与在完全随机的基础上去寻找相比较而言，你找到新矿藏的机会就会大大增加。

这种寻找满足问题所有条件的结果集合的穷举法叫做**蒙特卡罗**(Monte Carlo)* 反演，它通过概率法则来确保以下目标的实现：如果你以设计很好的搜索方式试验足够多的可能解决办法，很有可能你会找到许多满足所有要求的解决办法。实际上，蒙特卡罗方法只是向前模型的一个重复应用。总的来说，这一方法定义了问题的可能解决办法和不可能解决办法，而且对于可能的办法，发展了一些针对每一可能性的概率估计。真正的科学问题，比我们刚才研究的简单例子要复杂得多，有时要对千万个可能办法进行试验来发现那些满足我们所加的全部条件的相对较少的解决办法。

---

* 蒙特卡罗方法，也称统计模拟方法，是20世纪40年代中期由于科学技术的发展和电子计算机的发明，而被提出的一种以概率统计理论为指导的一类非常重要的数值计算方法。是指使用随机数(或更常见的伪随机数)来解决很多计算问题的方法。与它对应的是确定性算法。——译者

蒙特卡罗分析的一个例子和解释近来沿着格陵兰岛大冰原钻的深钻孔观察到的温度分布线图相关。[3]地质学家测量地球内部的温度有很多原因。一个原因是确定有多少热量从较深的内部流向地球表面，此热量是随位置不同而不同，并且和地形构造的稳定性相关。第二个原因是为了重构地球表面温度的历史，以便及时对后面地球的波动性气候有所了解。因为气候的变化发生在地表，地表下的物质将感知这一变化，因此后者的努力是可能的。结果，几个世纪以来延长的冷却过程将影响约1000英尺深的岩石温度(或者冰的温度，就格陵兰岛钻孔事例而言)。一个冰期将让3000英尺以下的岩石受到冰冻。

任何地上凿洞内部的温度分布线图是由向上流动的深层热流和向下传播的气候波动组成。对在格陵兰岛冰洞内测量的温度[4]进行蒙特卡罗分析，试验超过300万种深层热流和表面气候历史的混合，这些试验中仅有2000个能够充分解释实际所测量的温度。这2000个解决办法中的绝大多数显示了共同的特征，虽然它们没有以100%的确定性告诉你什么确已发生，但是对于格陵兰岛所在位置的实际气候所经历的历史来说，它们定义了一个相当小的范围。

贝叶斯和地上凿洞

在科学观察的世界中，科学家经常面对观察和不相配合的解释性假说。这就好像努力把一枚方钉钉到一个圆洞里。方钉与圆洞不是很适合，但这是钉子出问题或者洞出问题了吗？对于每一个而言，做什么样的调整能够让它们兼容？一个更大的洞或一枚更小的钉子可能是一种解决办法。一个不很圆的洞和一枚不很方的钉子，可能也很好地适应。“反演”的一种类型，以18世纪英国牧师贝叶斯

(Thomas Bayes)* 命名的贝叶斯分析，给了各种场景下概率的数量分析，这种场景的条件是一个人对此事例中的各种因素有不同程度的把握。[5]如果我们有理由相信钉子的形状并不比洞的形状容易理解，贝叶斯分析将估计这两者的最可能的形状，与我们先前关于钉子和洞的知识以及对它们为什么不匹配的感觉是一致的。贝叶斯方法与更加传统的概率估计方法相比，在概念化方面有相当的不同，原因在于它允许科学家进行判断，这些判断是对所研究问题中的所有因素的了解程度和把握程度进行表述。

概率估计中更传统的方法被称作*频率论者*(frequentist)方法，此方法通过一个事件发生的频率来进行推断。在经典的硬币投掷问题中，我们要问在任何给定的投掷中，我们观察到硬币正面或硬币反面的概率是多少。频率论者会投掷许多次硬币，把正面朝上和反面朝上的次数列成表，从正面朝上和反面朝上的频率中计算概率。然而贝叶斯分析，也许对此问题的处理相当不同。显然，只有两种可能的结果，通过求助于以前关于在空中翻转的金属圆盘的物理学知识，试验者能够确定这里没有物理理由让硬币的正面优先朝下或朝上。这种先前的知识暗示了硬币正面或反面具有相等的概率，不需要在试验上多次投掷硬币就能得出这一结论。

当我们在选举之前检查民意测验的结果时，我们已经讨论了一个频率分析的例子。“候选人史密斯的总得票数落入 38%～46%的范围有 95%的概率”的声明意味着，如果投票重复 100 次，候选人史密斯的总得票数将有 95 次落在此范围内。频率论者视概率为发生的频

---

* 贝叶斯(1701～1761)，英国牧师、数学家。曾经试图利用概率论证明上帝的存在。写了题为“机遇理论中一个问题的解”的文章，成为贝叶斯学派的奠基石。他的贝叶斯分析主要是采用过去事件的知识预测未来事件。——译者

率，他根据“如果一个人多次重复一个试验，会得到相同结果”的可能性来评估一个假说。贝叶斯分析家并不思考重复一个试验而是提出这样的问题：“这种单一的观察同多少种假说一致？”就像在钉子和洞的例子中，它将确定对于钉子和洞做什么样的调整能够使试验和它的解释前后一致。

在我自己的地质学工作中，我运用贝叶斯分析来推断最近地球表面的气候变化。就像我在“作者自我介绍”中简单提到的，我的研究集中在测量地球的温度。这种测量是通过将地球岩石凿洞到地球表面以下几百米的深度来完成的。我们沿洞的方向放置一支灵敏的温度计，大约每隔 10 米略微停一下来读取温度。这样就得到了一系列温度测量值，通过这些测量值总的来说能够画出沿凿洞方向的温度变化线图。在以上关于蒙特卡罗分析方法的讨论中，我提到了一个在格陵兰岛测得的温度变化线图。我和国际同行已经从各大洲的 700 多个洞里收集了相似的图线，并且已经分析这些图线来重建这些洞所在地在历史上经历的表面温度。

在这种背景下，一个贝叶斯分析进程如下：我们收集了一些在不同深度下温度的测量值，常常也能在钻孔期间，获得洞所在深度处的岩石样本。我们能够测量岩石的热传输特性来确定它们传导热量的能力强弱和速度快慢。这一物理特性在估计地球表面的热量向下传输所花费的时间以及影响不同深度处岩石的温度方面扮演着重要的角色。当然，回想起来，温度测量和岩石特性的测量都有着许多的不确定性，这些不确定性是与测量岩石热传导能力所用到的温度计和其他仪器相关联的。

下边我们做一个有根据的猜测(即提出一个初始模型)，此猜测是在我们知道岩石的热传输特性的情况下，什么样的表面温度历史是造

成我们观察到的地表下温度的原因。这一初始模型是验证它与观察的兼容性的一个假设。此兼容性试验对于地表下温度以及计算的温度与观察到的温度的兼容性来说，需要对最初猜测的结果进行计算。观察到的温度和计算的温度永远不会完全相同，因此计算机必须寻找这些或大或小的不匹配发生的原因。是沿着地下凿洞方向的温度测量和深度测量方面的不确定性引起的还是确定岩石特性方面的不确定性引起的？是因为岩石取样不充分，从而我们没有看到岩石热传输特性中的一些重要变化吗？是因为我们对温度历史的最初猜测不合适吗？是因为我们用于计算向地下传播的地球表面的温度变化所用到的热传输理论不充分或不完整吗？

贝叶斯反演会把所有这些因素加以考虑，在这些不匹配当中，对于有多少是测量错误引起、多少来自不充分的采样、多少来自热传输机制的不全面描述、多少来自对历史的粗劣猜想，它都会做出最好的估计。接着它会建议这些数量中的每一个必须如何改变才能达到相互一致。人类解释者将告诉计算机有多少可调节的余地。而且正是在后者的努力中，即估计在测量中、理论上和在对地球表面温度历史的最初估计中有多少可以调整，科学的判断和经验发挥着决定性的作用。

这里有两种极端情况经常由于过分严格被迫摒弃。第一种情况是，科学家坚持他（或她）关于历史的最初猜想是绝对正确的，而且如果地表下的温度没有同它相一致，那么温度测量就有一些错误。这是我早先提及的一个严格例子。“我的思想形成了——如果观察和我的模型不一致，那么就要在测量中寻找错误来解释这种不一致。”第二种情况的严格恰恰与第一种相反，“观察是绝对正确的，如果模型计算与这些观察不一致，那么模型计算所依据的模型或理论需

要被修正。”

当然，在实践中，我们知道在测量中具有不确定性，而且我们采用的热传输理论只是对于热通过岩石如何运动的一个近似描述。我们肯定不相信我们会预先完全知道表面温度的历史。我们考虑了一定范围的灵活性，计算机在其范围内能够对这一问题的每一部分做调节，而且会告诉我们这个拼图板的每一部分是如何最好地组合在一起的。

在关于岩石温度的解释中，我们对气候历史最初做了一个保守的猜测，认为最初根本没有气候变化。这被称作“零假设”（null hypothesis）。在提出零假设作为最初猜测的同时，我们也告诉计算机如果温度观察沿那个方向推动，我们将偏离保守的假设，但是无论发生多大的调整，我们都要加以限制。下一步计算机调查地表下的温度，来看它们在为温度测量所指定的不确定范围内是否与这种假设相一致。如果在此范围内不能达到一致，那么计算机转而会将气候历史的零假设调整到同地表下的温度更兼容的程度。

在这一过程的结尾，我们会看到如何修正模型以便与温度取得一致的贝叶斯估计，同时如何修正温度以便与修正的模型一致。这是一个高度流动的过程，在这里问题的每一部分，观察、理论以及解释，对于一些调整来说是公平的游戏。允许调整的范围是由“我们关于每一部分了解的程度”的最好判断和“哪一部分最容易弯曲或者最不容易弯曲”的最好估计支配的。如果计算机到达一个绝境，即在我们对每一部分施加的限制内，它不能在整体(pot pourri)之外再编造一个连贯的故事，那么它将发出抗议性的警告，称不能再帮助我们了。

不确定性与这一复杂菜谱的每一原料相混合，它也渗透在烹饪的说明中。但是贝叶斯反演过程不平衡地分配这种不确定性，最终我

们更好地理解了存在于我们观察、理论、“回答”和重建历史中的粗略性。如果我们进一步拿烹饪做类比，以下的事情是完全可能的：那就是当我们离开厨房时，如果蛋糕已经坏了或烧焦了，我们得到的不仅是一个不完美的产品，也有对原料是坏的、菜谱是错的、火炉是失灵的或者烹饪没有按说明进行的概率估计。

从世界各地的700多个地下凿洞的温度变化线图中，对于气候变化，我们了解到了什么呢？它们讲述了一个不寻常的故事，它是不依赖于表面温度的测量，这些测量是通过气象学观察做出的，或是通过用于测量海洋表面温度的海上浮标和轮船做出的。地下温度显示岩石在过去的5个世纪温度已经上升了1℃(差不多2℉)。但是上述升温中，至少有一半是在20世纪单独发生的，另外有30%发生在19世纪。这一升温，就像在岩石中看到的，和从地表温度的观察所重构的升温记录完全一致。因此根据这些来自地面下的数据，证实了历史的表面观察中所揭示的关于全球变暖的迹象。

相关性和因果关系

在重建过去中，我们不仅想知道发生了什么事情，而且如果可能，还想知道为什么这件事情会发生。例如，我们知道恐龙是在中生代末期灭绝的，但我们也想知道是什么因素导致这种灭绝。为了帮助回答因果关系的问题，科学家经常试图了解是否某一因素以一种模式运行，这一模式暗示它可能与科学家努力理解的事件相关。在气候变化领域，“过去一个世纪中地球的平均温度有明显变暖的趋势”这一观念已经很好地建立起来甚至被很好地接受。但是是什么正在引起气候变化？是来自太阳日益增加的辐射吗？是由于燃烧化石燃料大气中增加的温室气体吗？是因为火山活动吗？或者可能是

这些效应的一些组合引起的吗?

气候科学家用来表述这些问题的一种方法叫做**相关分析**(correlation analysis)。在与气温重建相同的时间间隔内,通过重建大气中温室气体浓度的历史,我们能看到在20世纪气温上升的趋势伴随着大气中温室气体浓度上升的趋势。每一个行为看起来都有随时间变化相类似的模式,我们通过以下方式描述这种相似性:地球表面的平均温度和大气中温室气体浓度是正相关的。这一相关性通过计算一种被称为相关系数的东西能够建立在定量基础上。

在没有检查这一计算复杂性的情况下,让我特别指出这一相关系数的一些特性:如果地球表面温度的每一个小的改变与温室气体浓度的相应的成比例改变,无论上升部分还是下降部分,都是并行的,我们计算出一个1.0的相关系数,称这两个是完全相关。如果一个的上升伴随着另一个的下降,反之亦然,我们计算出一个-1.0的相关系数,称其为完全负相关。如果一个的上升和下降看起来总是和另一个的上升和下降不相关,我们计算其相关系数接近于0。我们将说这两个是不相关的或可能不相关的。

然而相关性,与因果关系是不相同的。事情能够很好的相关,暗示着它们通过一些过程是相关的,但是过程本身并没有确定。所有统计学老师喜爱的经典例子是关于一个对火灾着迷的人的故事。他跑到每一个报告火灾的地点,并及时指出市政府消防署的消防车已经到了。他对于火灾的逐日逐年的日志记录,证实了火灾发生和消防车出现之间的强相关性。从这一系列观察中人们能不考虑"消防车是火灾发生的原因"的推断吗?单单是日志只能证明这种相关性;关于"火灾和消防车哪一个在先"的更仔细的观察,对于确定谁是原因谁是结果是非常必要的。

另一个众所周知的相关性是疟疾在人均国民生产总值较低的国家有着较高的发生率。[6]简而言之，疟疾和贫困在许多地方是手拉手前进的。不像地理学中关于火灾和消防车的相关性的虚构的例子，在那个例子中确定哪一个先来可能导致对原因和结果的理解。然而，疟疾和贫困相辅相成，每一个都是原因，每一个也都是结果。疟疾通过影响生产力、个人生产率、旷工、卫生保健花费、个人资金积累和国家资金积累等方面确实妨碍经济发展。但是落后经济由于不能提供公民的和公共的健康基础设施来消除蚊子繁殖区域，也不能提供医疗预防、治疗以及医疗教育，这样确实加速疟疾流行。从疟疾和贫困的相关性中，我们了解到，越富裕的国家，疟疾可能就越少发生；但是相反的情况就糟糕了，越贫穷的国家，疟疾发生的情况就越多。

有关“温室气体在气候改变中扮演的角色”的问题也进行了类似的研究。大气中温度和温室气体浓度的最长记录来自对南极洲长形冰核的分析。南极洲是高海拔和极地环境，那里的雪经年累积，每年的降雪都会被第二年的雪覆盖。柔软的雪几年之后逐渐被压缩成流动冰。一些进入雪中的空气被隔离在冰中的小气泡中。气泡中的空气是在降雪和再结晶之间间隙内的大气样本，分析此气体可以揭示大气中的温室气体在被封锁和保存之前丰富度如何。

南极洲极地高原钻探差不多已经穿透了 4 公里的雪和冰(12 000 多英尺)，这些雪和冰是在过去的 420 000 年积累而成的。雪沉淀处的温度能够通过每一冰层的水分子中氢同位素比例(重氢比氢)来确定。对气泡中空气的化学分析已经显示温度的变化和大气中温室气体浓度的变化是十分相关的：当温室气体浓度高时，温度也高；类似的，当一个低时另一个也低。但是此相关性并没有告诉我们是温室气体的改变引起温度的改变，还是温度的改变导致温室气体的改变，

就像火灾和消防车一样，一个详细的变化年表证明也许是有帮助的。一个是在另一个之前发生，或者反之亦然？在南极洲冰核包含几乎50万年历史的情况下，当某人努力在时间上往后看得越远，他对时间的分辨能力就减小了，问题仍然不能被明确解决。也许温度先增长，然后温室气体紧紧跟上来，或者沿相反方向变化。任一解释都需要让两者耦合到一起的一个机制。

这种耦合机制会是什么呢？几乎可以肯定，在全球气候系统中，它是第三方面的因素——地球上的生命。生命进入方程式是因为它是地球上碳循环的一部分。碳是所有生命形式（陆地上或海洋中大大小小的植物或动物）的组成部分。碳从先前的生物系统中进入大气有两种方式，一种是诸如植物死亡、腐烂；另一种是燃烧诸如煤、石油和天然气之类以碳为基础的化石燃料。燃烧和腐烂都涉及碳与氧化学组合产生二氧化碳气体。当大气中二氧化碳浓度增加时，海洋会吸收一些二氧化碳气体。

气候、碳和生命复杂地缠绕在一起，而且总是这样。其中一个变化会导致其他变化发生连锁反应。温度和大气中温室气体浓度强相关的重要意义并不是“小鸡和鸡蛋，谁在先”的问题，而是说明它们是相互强有力地联系着的。至于疟疾和贫困，50万年来的气候行进到一起的相关性，就像南极洲冰层记录的，是“当一个变化时，另一个也变化”的一堂课。在当前背景下，当大气中的温室气体与前工业时代水平相比增长了30%多时，我们应当从过去吸取的教训是能够预料到温室气体的增加伴随着温度的增加。

意识形态的扭曲

重建人类历史有着与重建自然历史同样的困难：不完整信息、不

准确信息和冲突的信息。但有时有另外一个因素会增加额外的扭曲。偶尔，通过有目的地缩减或删除所获得的信息，人类历史会被降低。当历史被迫通过意识形态的过滤，以一种“希望得到的国家形象”的净化形式出现时，这样的扭曲就产生了。在日本，一些中学历史教科书仍然忽略报告第二次世界大战中日本军队在韩国、中国、菲律宾和东南亚的其他地方等占领区所犯下的暴行。强迫妇女为日本军人提供性服务、细菌战的使用、残酷地对待占领区平民，这些是编写日本历史教科书的那些人更喜欢忽略的历史事实。

类似的，苏联在斯大林(Josef Stalin)统治下的历史许多年来被渴望忽略布尔什维克统治缺点的政府当局净化了。只有到了1989年铁板一块的一党制政府解体之后，俄国人在学校中才开始了解到残酷是1917年大革命以来的30年中他们政府统治的特征。而且对于20世纪的绝大部分时间来说，美国教科书也大大忽略了它在西部扩张，对国家土地的吞并，强迫美洲印第安人重新迁居等不光彩的方面。事实上，历史遭受扭曲不仅仅是由于来自一个不完整的记录，当历史学家应用了一个选择的过滤器时，也会发生扭曲。当科学被迫通过一个意识形态的过滤器时，科学也会遭受扭曲。如果生物学教师必须遵守神创论者的意识形态，就像在田纳西州发生的著名的斯科普斯(John Scopes)案件*一样，对生物进化的理解一定会遭受困难。

---

* 案子的核心是高中教师约翰·斯科普斯在公立学校课堂上讲授达尔文进化论，因而违反了田纳西州的法律——这条法律直到1967年才被废除。这一著名案件引起国际社会注意，名律师克伦斯·达娄(Clarence Darrow)为斯科普斯辩护。尽管斯科普斯被定罪并处以罚款(但从未被强制要求支付)，大多数人认为，田纳西州从此案中所得的丑名与嘲弄，导致的结果是科学与进化论教学的胜利。然而此后的事实并非如此。在案子进行期间和结束之后，人们的情绪都很激动。多数出版商采取了在高中课本中忽略进化论的策略，为的是避免书的销路被争论所影响。一代又一代的美国学生并未接触到进化论这条生物科学最重要的基本原则。——译者

过去总是融合着未来。今天是你昨天担忧的明天。[7]对于继续越过当前边界的过程来说，对过去的理解是预测未来的钥匙。但是只有当过程和环境不随时间改变时，过去才是未来可靠的钥匙。然而地质学和历史告诉我们，静止的世界只是幻觉。未来是和过去相关的，它建立在过去之上，但它很少是过去的精确重复。在下一章，我们探索预测未来的不确定性和风险。

注　释：

1. *Science*, 7, February 1890.
2. Dated 22 August 2000; www. ntsb. gov.
3. Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., et al., Past temperatures directly from the Greenland ice sheet. *Science* 282, vol. pp. 268～271, 1998.
4. 在这种情况下，冰传导热的能力就像岩石或任何其他固体物质。
5. 一个简短的对贝叶斯方法的描述可参见 David Malakoff in *Science*, vol. 286, pp. 1460～1464, 1999。
6. Sachs, J. and Malaney, P., The economic and social burden of malaria, *Nature*. vol. 415, p. 680, 2002.
7. 这一精巧的措辞归属于 Jerry Longan。

第十章　预 测 未 来

想得太远是个错误，每时每刻只能抓住命运的一环。

——温斯顿·丘吉尔

预测未来……前景是多么的诱人。预测未来对于许多人来讲已经成为一种职业。仅需要一点努力，我们就能发现算命者、巫师、看手相者、占星学家、神秘主义者、预言家、心灵学家和许多其他为了赚钱乐于揭示未来的人。但是我们都应该对任何神秘的通过特殊途径知晓未来的从业人员多些怀疑。甚至是使用不太神秘工具的经济学家和专业人员——气候学家、保险精算师和养老金管理者——发现遥远的未来是很模糊的，通向未来之路是坑坑洼洼、崎岖不平的。本章最重要的主题是：未来是一个移动的目标，预测它的特征总是困难的，而且试图向前看得越远，预测未来就越加困难。处理这种不确定性，我的观点是发展长期的眼光，制定向前推进的计划。但是沿着这条路线，要准备进行许多方向修正，因为未来的展开与你已经预言的有相当的不同。

我们都听说过对于未来唯一确定的事情是死亡和税收。这个惹人喜爱的格言确实说明了未来大多充满着不确定性。然而，不确定性是不统一的，未来的一些方面比其他方面清晰。除了死亡和税收之外，我们应当确信我们将不能赢得彩票和太阳明天将要升起。[1]未来的绝大多数其他方面位于不确定性光谱中的灰色区域。

事情的真相是，预测未来尤其是非常遥远的未来的确是非常困难的。许多不久前作出的预言今天被认为是偏离目标太远，甚至被认为是非常可笑的。让我们思考一下几个来自当时受人尊敬的专业人员的评论吧：[2]

> 飞机是有趣的玩具，但不具有军事价值。
>
> ——费迪南德·福煦(Ferdinand Foch)，
>
> 指挥官和战略学教授，法国军事学院，1907～1911

> 股票已经到达看起来永远是高原的地方。
>
> ——欧文·费希尔(Irving Fisher)，
>
> 经济学教授，耶鲁大学，1929

> 我认为世界市场也许只需要5台计算机。
>
> ——托马斯·沃森(Thomas Watson)，
>
> IBM董事长，1943

> 我们不喜欢他们的声音，吉他音乐即将灭亡。
>
> ——麦加唱片公司(Decca Recording Company)，
>
> 反对披头士，1962

也许我们会笑话这些例子，但是同我们对未来的思考相比，这些例子是对未来作长期预测的典型代表。未来的领域撒播着这样的偶然性，结果让人不能更加清晰地看见未来。然而，对一些预言作中期的方向修正，就能让它们沿未来的浪潮行进。20世纪历史显示，IBM发现将有一个比沃森预言的更大的大型计算机市场，而且对于家庭能有效使用的小型个人计算机来说，有一个灿烂的未来。IBM在20世纪80年代引入个人计算机就是对沃森预言做了一个很大程度的方向修正。

预卜未来

一个人可以从很多方面发现有关未来的消息，如在肯塔基的德比或者皇家阿斯科特，哪匹马将会获胜？纳斯达克(Nasdaq)或(英国)金融时报-证券交易所(FTSE)的哪只股票会成为领头羊？这些消息来自那些人们，他们专门研究过去那些成为获胜者的马或者股票取胜的原因。有些马在短跑中表现更好，有些马擅长在潮湿跑道上奔跑，有些马在不太拥挤的竞争中表现突出，有些马喜欢在凉爽的日子表演，还有一些马在有可靠职业赛马骑师的情况下发挥更好。同样股票分析家也开发了评估企业的方法：公司产品质量或服务质量、产品需求或服务需求、管理力度、所欠债务负担、员工的忠诚度、竞争力。当我们购买关于赛马或者股票的赌注时，在某种程度上我们理解在这些决定中具有风险和不确定因素。而且通常当这些预测不成功时，我们会接受冒险的结果，也许会寻找关于这些事件的其他建议。事实是不同的人要选择不同的马，这是古谚语“那就是决定赛马的东西”的要点。下在每一条赛马跑道的赌注或者对特定股票的每一次个人购入或抛出，都反映了对某一未来活动的可能性评估。

我们有一个充满活力的赛马行业和生气勃勃的股票市场，这一情况证明了人们在不确定面前面对着不确定的未来，运用最好判断和采取措施是一种自发行动。

另一个对待近期未来的熟悉事例是天气预报。在这一活动中，工作中的专家研究大气物理学和流体力学，他们运用大型快速计算机收集每小时气象学的观察和轨道卫星的信息流（information streams)。诸如夜间天气预报，甚至5天内天气预报的预测已经成熟到一个可靠阶段，借此它们在旅游计划、农业管理、遏制森林火灾和战争中扮演着重要角色。但是，对于全球金融市场作类似的5天比较有把握的准确预报甚至不是地平线之外遥远的可能。气象学预报和经济预报之间的差别只是反映了这一事实：我们对气象学的理解远比对经济的理解要透彻。

以上的每个例子，赛马、天气预报和投资都涉及不确定性在不同时间尺度内的延伸。赛马的获胜者在比赛开始之后的几分钟内便可知道，当然很快就能够对某人先前的信息和建议进行评估。气象学家一小时接一小时地使用新数据更新他们的计算模型，并对每日的预报进行修正。投资的成功性也许对每天、每月甚至每年来说不是很明显，但是当未来展开时，如果早期的投资路线已经引起危险，投资者必须准备调整他们的资本组合以便回到充满前途的道路上。

对当地气象条件做一年或两年的长期预测实际上达不到，因为一年年影响这些气候条件的许多地区因素，例如在太平洋赤道厄尔尼诺南方涛动或大西洋北部的高纬度地区的北大西洋振动，它们本身在持续时间内是变化的，很难进行预测。此外，这些区域性的特征通过远程并置对比连接于全球系统。太平洋地区厄尔尼诺现象的强弱一年后才会在津巴布韦的降雨中被感知。然而，对10年后全球平均条

件的估计并不困难，因为对这种未来的预测要求更少，在全球平均预测中区域性细节和短期的时间细节“最终会得到平衡”。在某种意义上，某人正在寻找的仅仅是未来的一个大的轮廓，而不是每一个区域、每一天的细节。人们问的是气候问题，而不是天气问题。

经济上也有对于过程的长期预测。我们听到确信的预测：经济衰退在经济增长恢复之前要持续10个月，或者利率在来年会降低，因为国家发展经济的形势发生了变化。随着时间的推移，这些预测成功与否常常和随机波动难以辨别。对比来看，长期的气候预报和对10年之后平均条件的估计，在类似时间尺度内可能比经济预测更牢靠。

短期经济预测和长期经济预测的缺点根本上来自于对影响国家和全球经济的复杂过程和因素的知识和理解不够全面。对比起来，天气预报是容易的。大气科学家实际上对于支配大气短期行为的物理过程理解得相当好，所以随着一系列气象站和卫星把对太阳射线、大气温度和大气压力的观察输入大型快速计算机中，我们能够得到合理准确的天气预报。但是对于支配股票市场行为的过程，我们能够说些什么呢？在复杂的金融领域，是什么经济规律或行为法则统治着每天的波动和长期的趋势？这些比物理学规律更加模糊不清，因此在国家或全球经济的王国，明天、下周、明年将要发生什么？对此问题有着更广泛的观点和大量的不确定性。

每一位市场分析家都确认他或她相信成为支配标准的东西，而且基于这些标准来做个人对于未来的计算。实际上每一个人都创建了用于预测未来的市场行为的个人模型。每次预测都是我们称之为“由模型决定”的东西，即结果由那个人在作预测时对于影响市场进行的过程和因素所理解和相信的东西决定。这些分析家已经正确地确认了相关的重要标准，在这个意义上，对未来的预测会或多或少有

些准确。这种有关不确定的“智慧”等待出售，投资者必须运用自己关于“哪一观点更有价值”的判断或直觉进行选择。一些人甚至从许多观点推算出一致意见或者对未来作出“平均”预测，好像每个观点都是对未来经济情况的不准确测量或估计。

未来市场水平的巨大不确定性基本上来自于控制市场的过程和因素的不确定性。这种不确定性，我早先称之为手头问题的概念化引起的不确定性。就市场行为而言，这种概念化的不确定性远远超过在就业水平、清单、工厂产量以及诸如此类的经济因素测量中的任何不确定性。

算错

由于有事后之明的好处，我们知道预测未来经常是不相关的，不可预料的危险并不少见。不久以前在我生活的地区有一家重要的报纸刊登了这样的标题：“吞没高速公路的增长比预测得还要快”。[3]标题下的文章指出对最近刚完成的底特律北部的州际高速公路系统链环的交通容量的预测是非常不准确的，当前的交通容量甚至已经比对未来10年的预测要大50%。这未曾预料到的交通容量归因于不可预料的经济增长和住宅增长。在讨论如何估计交通容量时，地区政府委员会交通计划的主管说：“这种假设没有必要转变成现实，目标从来不会停止移动。一个变动公司(地点)的董事会决议能够改变事物。”

显然，对未来几十年交通模式和交通容量的估计也包含着不确定性。那些尽力处理拥堵的计划者是愿意从过去经验中学习，还是只是愿意重复过去的错误？本文讨论的特别的解决办法包括增加更多的“开—关斜坡”和更多具体的车道。但是对于未来某种程度的拥堵，公众会意识到此问题的基本概念化可能是错误的。问题也许应

被认为如何把人们从他们所在的地方移到他们想去的地方，而不是如何移动汽车。一旦计划者和决策者从把汽车作为唯一交通方式的思想中解脱出来，就有可能考虑其他的可能性。在华盛顿地区，30 年前拥堵达到一个阶段，在那里地铁交通系统成为有吸引力的交通替代品。今天，华盛顿的地铁是每天进出城市的千百万人首选的交通方式。世界上绝大多数大城市，如伦敦、巴黎、莫斯科，很久以前人们就认识到需要地铁系统以到达城市中心工作或购物。

被称作长期资本管理公司(LTCM)的华尔街公司开发的关于安全和金融市场的数学模型有一个作为其中心特征的轻率走进布雷区的故事。这个模型从 1994 年到 1998 年作为制定投资策略的概念和技术基础一直服务于长期资本管理公司。LTCM 已经建立了一个绝对可靠的形象，部分原因是它的建立者包括两名最近的诺贝尔经济学奖获得者，并且它对最初几年的投资产生了显著的回报。LTCM 从富有的投资者那里吸引了非常大数目的资金。它的基本的投资策略是通过复杂的财产计算确定投资产品之间一个很小的价格差异，这些产品在平衡市场内价格本应是相同的。LTCM 会利用这些同平衡市场暂时小的偏离进行大量交易。

这一策略，尽管听起来很有吸引力，但是当一个基本模型假设在现实中不能具体化时，它就摇摇欲坠了。缺点是什么？LTCM 的计算机模型设想总有一个市场：当你想卖时，就有别人会买。但是在 1998 年 10 月，由“8 月中旬国际上债券掌控在俄罗斯手中”这一事件引起的恐慌中，突然间很少有购买者让 LTCM 能够完成成对的交易。LTCM 已经买了债券但是发现卖的机会已经丧失。此外，LTCM 必须卖掉的许多债券是用借的钱购买的，而且贷款银行要求还款。为了偿还贷款，LTCM 被迫以打折的价格来清算其他的投

资。数十亿美元被投资者损失掉了，令人担忧的是LTCM没有能力偿还它的债务，这使贷款银行、其他金融机构和联邦政府作出了营救努力。这个营救团队担心LTCM的失败很快会导致银行倒闭，随后的金融波动可能会变成一个经济海啸，出现无法想象的后果。从中获得的教训有许多，的确，损失惨重的个人投资者有他们自己长长的清单。但是对于更广大的民众来说，一个重要的教训是银行应更加关注它们贷款的脆弱性。私人投资公司能够逐渐削弱主要银行的稳定性，这一现象对于贷款官员和那些制定支配高杠杆效应投资规则的人来说，是催人觉醒的呼喊。

就在LTCM面临崩溃时，另一个布雷区不经意地被展开了。1998年1月，在广泛的商讨之后，加利福尼亚州重建了它的电力工业，并取消了对它的管制。新的结构将发电从电力输送中分离了出来，这实际上打破了由州直接管制整个电力工业的历史结构。就在这一重构生效之前，一位新法令的制定者乐观地预言："将有一个全新的制度，电能将在一个开放的市场上进行交易……从现在起的5年内，电能的零售价将是现在的一半。"3年后，他说得既对又不对。是的，有了一个全新的制度，电能在一个开放的市场上进行贸易，但这个市场却是一个混乱的市场。在那些已经取消管制的地区，[4]零售价格不但没有减半，反而又涨起来了。批发价格已经暴涨到了极高水平，发电厂商牟取的暴利据说是猖獗的。加州州长和立法机构制定紧急临时法令来防止重建工业的传输和发送部分的破产，它们正受到发电厂和不满的顾客的严重挤压。但是就在电力不足结束之前，一家主要的电力贸易公司和加州最大的公用事业公司都破产了。此事件的尘埃尚未落定，就起草了新的法律来纠正最初计划中的缺陷。确实需要一个中期的方向修正。

当其他州仔细考虑或者开始规划公用事业的重构计划时，加州事件提醒人们要吸取一些教训。取消对公用事业的管制处于整个过程的初期阶段，犯错误也不足为奇。但是生手可从他们的错误中学习，因此我们也希望消费者、供应者以及立法者能够从加州的崩溃中受益。我们也许不能从任何有益的程度对未来进行确定的预测，但是我们能够认识到什么时候未来将以有害的方式展开，并进行必要的方向修正，让电力工业之舟驶入更加平静的水域。

不足还是过量？

不是所有的未来都要通过布雷区。在我能够回忆起的岁月里，用于支付美国联邦政府支出的税收收入对于国会和总统认为必要的工程和计划来说看起来是不够的。因此预算不足成为多年不变的一个特征，常常需要通过政府借款来资助，国家债务也是年年攀升。在刚刚过去的20世纪90年代早期，全国所有顶尖的经济学家都在预言，预算不足将持续增长；全国所有一流的政治家奋力创建预算平衡，如果这方面能够做到并且保持下去，将能保持预算稳定，但不会将国家的债务减少很多。

到了21世纪的第一年年初，已经驳回了所有这些令人沮丧的预言。在20世纪的最后几年里，美国经济的非凡成功将税收收入提高到空前的水平，国库充盈。接下来的谈论转到如何处理巨大的过剩以及国家偿付债务的过程有多快的话题。2000年6月末，估计财政收入盈余(包括社会保障和非社会保障组成两方面)在21世纪前10年几乎达到4.2万亿美元，[5]已经超过3.5万亿美元的全部国家债务。但是特别要说明的一点是，至少在经济预测事务中有多少不确定性的意义上，关于10年间盈余的估计与仅仅5个月前做出的估计相比，

就超出了30%。而且到了2001年初，仅仅6个月之后，关于盈余的估计又一次增长到5.6万亿美元，[6]比11个月前做出的预测多出70%。然而，2001年年中，盈余的欢欣被全球经济的衰退以及实质上政治性激发的追溯的减税削弱了。接下来到了2001年9月11日，恐怖主义者袭击了纽约的世贸中心。对这一悲剧事件的反应包括用巨大的联邦基金重建纽约，补助航空运输业，对机场增设额外安全措施以及对阿富汗开战。2001年底，预测在接下来的3年中联邦预算会出现赤字。

除了恐怖分子的袭击令人吃惊之外，很明显建立国家经济模型也是十分困难的，因此在联邦收入的预测中有着巨大的不确定性。但是这种不确定性不允许采取简单的方式等待形势被澄清。必须对预算进行明确表达，必须做出花费决定，还得考虑修订税收法规，所有这一切都面临着不确定性。不论我们是否开始着手减少债务、减少税收或者对计划性花费重新优化次序、或者经历新的突然袭击，我们都必须承认，几年之后，由于条件变化国家也需要一个方向修正。这一观念被引入到国会关于减税的争议中，在那里它被称作触发条款(trigger clause)或安全阀门，用于在收入没有达到预期数目的情况下，停止减少税收。不过它并没有被采纳。

耗尽石油？

古尔德观察到[7]“关于未来的所有让人感到极度痛苦的问题几乎都聚焦在短期的扭动上面，而不是在较长尺度上的更广的图案上”。但是我们如何对也许延伸到我们寿命之外的不确定性进行反应和表达呢？不时有人提出这样的问题：“什么时候世界会用光石油？”对此问题的广泛回答是混乱多变的。一方面，我们听到乐观主义者

说，“永远不会用完”；另一方面，我们听到“这个世纪确实是石油为全球经济提供动力的最后时期”。然而，关于石油有一件事情是确定的。地球上石油产生的速度是在一个缓慢的地质时间尺度上进行，它的消费却发生在一个快速的人类时间尺度上。换句话说，自然不会以和我们抽取石油一样快的速度制造石油来营救我们。在地球漫长的历史时间内已经产生的任何东西是我们工作时不得不使用的资源。

当然，有许多因素，将影响汽油产品在为机动车辆和发电站提供燃料、为家庭和办公楼供暖以及为石油化工工业提供原材料方面所能持续的时间长短。来自石油的能源需求可能会受到诸如更好的绝缘结构、家用电器和车辆发动机中更高的燃料效率、更小更轻便的轿车、更好的公共交通选择，当然还有新开发的替代燃料的发展之类的能源储备措施的诱惑。煤炭和天然气在全球范围内可广泛得到，核能已经用于产生世界总电能 10%～15%。在一些地区，太阳能加热、太阳能发电、风、水力发电和地热能也作出了实质性贡献。自然有与许多石油的替代品相联系的环境问题(当然包括石油本身)，但仅仅是增加千瓦时的简单运算，也有充足的能源能够减少而且也会减少对石油的需求。

价格也会影响对石油的需求。在整个 20 世纪，石油由于低廉的价格一直是最好的燃料。按真实价格(非波动价)算，以石油驱动的燃料价格在 20 世纪实际上是下降的。当我在 20 世纪 50 年代刚开始驾车时，艾森豪威尔(Dwight Eisenhower)刚刚当选美国总统，汽油价格是每加仑 20 美分。油箱中 1 美元的油足够城镇人们在一个繁忙的星期六晚上使用了。今天的价格在每加仑 1 美元到 2 美元之间，半个世纪以来价格大约增长了 7 到 8 倍。你还能想起什么别的东西

增长得没有这么多呢？住宅，不论是购买的还是租用的，以大于10的系数增长。作为大学生，我每天用于购买食物的预算为1美元（普遍承认是相当简单的饮食），然而今天每天花10美元也没改善多少。然而，石油相对来说仍然是便宜的。几十年了，石油交易是3美元一桶，除了短期内受国际冲突的推动，它以低于30美元每桶的价格一直保持到今天。[8]

石油的价格比日常生活消费增长缓慢的原因在于它一直是充足的。随着需求的增长，石油的供应也能够增长，因此也调节了天然气的价格。大型国际石油公司提醒我们过去它们没有让我们失望而且也完全有能力支持我们进入将来。但是石油的供应会充足到适应现代工业经济日益增长的能量需求吗？对此问题的回答是"什么时候世界将耗尽石油"这个问题的核心所在。

对上述问题的回答中，部分差异来自于对石油的定义。对石油最普通的印象是它是位于地球表面岩石下的一种液体，能够被引导流动到油井中，从井里通过泵能够被抽送到地球表面。然而，除了液体石油之外，还有不能流动和被抽送的固体石油。固体石油，以油页岩(oil shale)*的形式或被锁定在砂石小孔中以黏性的沥青形式存在，其存量大大超过了液体石油但不容易利用。"可能永远不会耗尽石油"的宣言通常是来自于对包括液体石油和固体石油的整个石油资源的考虑。经常遗忘不被提及的是，抽取固体石油的花费要比相对容易的抽取液体石油的花费大很多。实际上，抽取固体石油的技术还处于初期阶段，固体石油与钻探和抽取相对简单的液体石油相

* 油页岩是水藻炭化后形成的，含灰分过多，多半不能自烯。也是一种化石燃料。在美洲大陆偏多。——译者

比，本身是更密集的能量。因此，尽管在地球上依旧存在着巨大数量的石油，但大多数不是将来也不会很快是我们能够容易掌握的。

就液体石油供应来说，我们能期待什么？与公众的认识相反，也与许多石油生产商告诉我们的相反，我们已经看到了这一资源的有限性。这一宣言背后的根本原因是我们正在抽取的石油比我们正在发现的石油更多。这是一个必然导致资源耗尽的方程。当从海中捕获的鱼比通过自然繁殖替代的鱼更多时，鱼作为一种资源会减少。当从地下蓄水层抽取用来灌溉的水比通过表面渗透所补充的水更多时，蓄水层作为一种资源会减少。石油没有什么不同。当前世界每年消耗大约 260 亿桶石油，而每年发现的新石油仅仅大约 60 亿桶。[9]当消耗的石油比能够发现的石油多时，我们是走在通往资源枯竭的路上。短缺背后的统计数字是非常清晰和明确的。

为了响应增长的需求，美国石油产量在整个 20 世纪逐年增长，于 1970 年左右达到顶峰。从那之后，开始下降，今天维持在 1970 年顶峰时大约 80%的水平。无论开采方面的探测和技术革新多么迅猛，石油产量还会继续下降。简单的事实是石油变得越来越难发现，也越难于抽出地面。很容易发现和生产石油的日子显然已经远离了我们。[10]

然而，那不是你从石油生产者及其在国会中的朋友那里听到的故事。他们告诉我们，如果他们不受探测和钻孔地点的环境限制，就会有充足的石油被发现和生产。[11]但是过去的 30 年美国石油产量下降的原因既不是来自专横的限制，也不是缺乏技术更新。相反的，恰恰是在地理探测扩展到大陆架和深海时，同时是对地面下的东西使用超级计算机、三维成像和可视化技术使得探测技艺显著提高时，产量下降发生了。被发现的石油的产量甚至在钻孔和抽取技术取得实

质性改善时也在下降。如果在新的地质环境中没有发现石油，也没有生产技术的进步，从1970年达到顶峰以来的下降将更为迅速。

当然，需求根本没有减少。在1970年国内生产的顶峰时期，美国人每天大约消耗1500万桶石油，国内产量能够满足超过四分之三的消耗量。但是到了今天，虽然国内产量持续下降，消耗量却达到每天2000万桶以上。这必然导致了对进口石油日益增长的依赖，以便适应这种短缺，从这个意义上，2001年美国日常石油消耗中有56%来自于进口，与之相比较，1970年仅为23%。随着时间推移，这种对外国石油的依赖只会增长。在这种背景下，小布什政府将把进口石油减少到仅占国内消费的50%，主要通过增加国内生产作为其能源政策的一个目标。为达到这一目标，需要将国内生产增长到20世纪80年代以来未曾见过的水平，这实际上是对生产长期下降趋势的颠倒，也将是对过去历史经验的挑战。这是一个一厢情愿的想法的最好例子。

在美国已经发生了的事情将会很快超过全球石油产量：我们将看到世界范围内的生产高峰出现在本世纪20年代早期。当此事在2010年到2020年之间的某个时间发生时，世界将会开始新的经济路线。到那时，我们将看到下降的生产赶不上增长的全球需求。大约一半的液体石油资源仍在地下，但是抽取起来越来越困难。结果是足够清晰的。价格将实质性增长，由于石油廉价充足的日子离我们远去，形势不利时对石油的需求(但不是对能源)可能会减小。固体石油在召唤，但是抽取这一资源的技术不可能确保从传统的液体石油进行一个平滑的过渡。最初依靠的燃料可能会转移到天然气，一种与石油有着相似生产和消耗历史的燃料，但是时间上被耽误了。替代的能源场景也许原则上能够确保一个从石油的平滑过渡，但是当大约

10年之内传统的石油生产开始全球性减少时，如果要拉紧松弛部分，开发替代品的速度就必须加快。

未来的气候

另一个不确定的未来，刚刚显现初步的轮廓。这个不确定的未来是和下面的观点相关的：地球表面的变暖是由于一系列自然的原因和人为的(即人类引起的)原因而发生的。在逐渐发展的争论中(在最后一章我将抓住机会详细回顾气候变化)，不确定性每一次都集中于这一主题的许多不同方面。一开始，为“地球正在变暖”的论断提供支持的数据曾因不充分和不准确受到挑战。后来这一辩论变为，已经承认是事实的地球变暖，是否就像历史、考古和地质记录所解释的“显著超出了自然气候上下浮动的范围”。当另外的研究显示20世纪温度的上升的确是异常的巨大和迅速时，这一争论又转到变暖的原因上。

由IPCC(气候变化政府间专门委员会，该组织在第五章首次被提到，在联合国和世界气象组织的领导下工作)提出的压倒性的科学意见一致承认两种因素的重要作用：一种是诸如来自太阳的不同辐射和来自火山喷发的偶然动荡之类的自然因素；二是诸如温室气体增长，尤其是燃烧化石燃料在大气中形成的物质和烟雾之类的人为因素。科学研究表明，在过去的1000年中自然因素统治气候波动大约到1750年或1800年。从那时起到1950年左右，人为因素增长到足以和自然因素竞争的地步，导致一种非常复杂的混合力量所引起的气候变化。然而，在20世纪后半期，人为力量大大超过自然因素，人类在地球变暖方面的烙印已非常明显。但是不确定性的阴云，一些真正的科学和一些由工业利益驱动所散播的烟幕的所谓科学，已经

迷惑了公众，这使得人们无论是理解过去还是预测未来，在接受全球气候变化的科学基础时都会非常谨慎。

预测下世纪地球的气候不是一项容易的任务。研究这一广泛问题的自然科学家和社会科学家承认许多复杂的和相互依赖的因素将影响气候的变化。首先，地球的人口增加了多少人，他们将住在哪儿？他们对生活渴望的标准是什么？什么新能源会为全球经济提供动力？什么时候将逐步采用新能源？如果有人假装知道以上所有问题的答案，那么他们要么是欺骗自己要么是试图欺骗你。实际上，这些问题中每一个都是有许多可能答案的高度复杂的问题。事实上，与理解地球气候系统运行的方式相比，这些问题有着更多的不确定性。气象学和气候学，就像许多物理学方程和流体力学方程所表达的要依赖大型计算机解决问题，它们远比社会和经济的趋势有一个更坚实的基础，它们提供给气候模型所需要的数据就可以计算一个世纪或更遥远未来的气候。

IPCC 的科学家提供了关于未来的一系列图景，这些图景也许是整个 21 世纪人口统计学发展和经济学发展的结果。其中的一个图景家族是这样描述这一世纪的：本世纪是经济快速增长、迅速引入新的高效技术、到世纪中期人口达到顶峰然后开始下降的时期。此图景家族的特征是：地区和增长的文化交流之间的经济趋同、真正的“全球化”的经济、模糊的文化差别和地区差别。在这一普通框架内，有多种方式取决于驱动全球经济的能源选择：高能化石燃料、高能非化石燃料或者所有可能获得的能源的各种混合。

与上述相反的图景家族是这样想象的：面向全球化的趋势是十分缓慢的，而且世界各地区的发展更加不均匀。生产模式仍然明显不同，一些地区出现巨大的人口增长而另一些地区保持稳定。经济发

展区域性集中，技术改变的扩散更加缓慢，能源使用仍然主要依靠化石资源。

从经济发展的每一个模式来看，预测人口曲线、进行能源选择都要通过大气中温室气体的浓度和伴随的全球平均温度增长的复杂模型进行运算。这些不应当被当作个人对21世纪气候演变的预报，而是定义的“随时间推移我们将会遇到的气候条件的可能范围”的图景的总效果。了解在各种不同假设下的可能结果的范围也许比对任一图景的细节进行诡辩更为有用。这些程序能够被设想为一种对结果的系统化和方法的探索，一系列赋予不同实验变量值的数值实验。IPCC对于21世纪的气候预报既可看作过程，也可看作产品。

比较和评估多种图景的策略不是预告未来气候的唯一方法。它是评估未来经济条件很常用的方法。常常能读到这样的报告，“16家投资银行的经济模型显示，下一年经济在2.3%到4.1%的范围内增长”或者“24位专门研究债券市场的分析家预测利率在下一年将减少1%到3%”。至于气候图景，感知由许多独立评估决定的经济结果的范围比过分强调一个特别模型的复杂性常常更为有用。

通过对多个图景的系统化评估来分析未来是在第九章中讨论的钱伯林的多种工作假说的一个变形。对未来做决定充满困难的原因如下：这些决定处在不确定性的包围之中，而且不确定性在一个有益的时间框架内不可能被解决。因此做决定必须要对许多选择进行评估，对未来不同选择的结果做评估。保持现状通常就是受到惯性统治的一个选择。这一选择经常受到那些对保持现状有既得利益的人和那些会受到偏离现状决定不利影响的人的拥护。在对包括保持现状和那些涉及显著变化的选择进行评估的过程中，我们必须理解引发各种选择的哲学基础和经济背景。在每一图景中谁将是获胜者和失

败者？我们如何权衡那些明显是从自身利益出发的各种论点？做决定时一个合理的延迟是否允许减少不确定性的新信息出现？我们能买用于减少不确定未来的不利后果的“保险”吗？所有这些问题如果能够被有效地处理，就能照亮不确定性花园中通过灌木丛的小道。如果我们偶尔被困在花园的迷宫里，那我们必须从一些隐蔽的小路和死胡同退出来。就像我在本章开始提到的，很难清晰地看到通往未来的道路。关于未来是很容易犯错误的，但是这不应该阻止我们穿越迷宫，当然得承认，一些中期的方向修正可能是必要的。

“多种工作假说”作为一个策略，强迫人们对未来进行一系列思考，即一系列可能的路线。它强迫人们保持思想开放，评估选择和考虑方向修正。保持思想开放和用新思路思考问题是只从一个思想观点考虑问题的对立面。意识形态在不确定性花园出现得很频繁：在第二章它出现在“不确定性的传播者”上下文中；在第七章出现在“概念化”和“在盒子之外思考”的上下文中；再一次出现是在第九章“意识形态过滤器是如何扭曲历史”的上下文中。受到意识形态的影响等同于将可能性图景的家族缩小到一条单一的道路。一个封闭的思维只能看到一条路，很可能最终演变为一种顽固的习惯思维。

在这一章中我试图阐明如下情况：处于许多条件下的未来不能被非常清晰地看到，一个人试图看得越远，未来就越加黑暗。方向修正对于在限制太多之前逃脱习惯思维是必要的。但是有时我们对于未来的视野不是受到幽暗的照明的阻碍。偶尔，未来的窗户整个是不透明的。这些是我们一方眼睛看不见的时候，是让我们完全惊奇的事件发生的时候，这样的事件发生在2001年的9月11日。像这

样的事件是下一章的主题。

注　释：

1．地球绕其自身的轴旋转的规律性甚至在莎士比亚(Shakespeare)看来也是非常明显的，他写道“……一定遵循昼夜更替这样的规律……”(*Hamlet*，Act Ⅰ，Scene Ⅲ)。

2．这4个引用语出自Smithsonian Institution 5年(2001～2005)巡回展“Yesterday's Tomorrows：Past Visions of the American Future”的一个长表中。

3．*The Detroit News*，2 July 2000．

4．在加利福尼亚，并不是所有地区都取消管制。一些市政当局，最典型的是洛杉矶，发电站仍然由市政当局拥有并进行管制。这些地区要保护它们不受这个州其他地区因取消管制而引发混乱的影响。

5．*New York Times*，27 June 2000．

6．*New York Times*，31 January 2001．

7．This View of Life，in *Natural History*，September 1998．

8．对于那些没有特殊技能的工人来说，其最低试用工资增加得很少，仅从1950年的75美分增加到2001年的3.75美元，只增加了5倍。

9．Walter Youngquist，*Geotimes*，pp．24～27，July 1998．

10．Kenneth Deffeyes的著作*Hubbert's Peak：The Impending World Oil Shortage*，Princeton University Press，Princeton，NJ，2001，285pp，其主题是石油生产的历史及其对未来的影响。

11．在回答记者关于美国石油和天然气的供应情况的问题时，代表美国白宫的一位重要人士回答：“我们有无限的供应，只是还没有发现。”

第十一章　完全出于意外

往最好处期望，朝最坏处打算，并且为意外做准备。

——丹尼斯·惠特利(Denis Wheatley)

不能预料的事件，我们称之为随机事件，如果在它发生时我们不知道是什么原因引起其发生的，那它必然是不确定性的来源。在个人层面上，我们可能熟悉许多这样的事件：汽车事故、家中发生火灾或盗窃、家庭顶梁柱(the family breadwinner)的突然死亡甚至更为悲惨的恐怖袭击。其实在地区甚至全球区域的基础上，也可能发生其他不幸，这些不幸具有与发生在个人层面上的不幸同样的突然性和灾难性。虽然我们也许不能通过直接的个人经验认识这些不幸，但我们也熟悉这样的事件：泛滥的洪水、大范围的停电、破坏性的地震、巨大的火山爆发。不论是个人层面还是更广泛的层面，所有这一切都是我们特别不想经历的事件。然而在任何指定的一天，我们在生活中可能会遇到上述的某个事件，我们知道这种可能性虽然很小，但

不是无限的小。相应地，我们开发了一些通用的策略来处理这些事件的后果。为了减轻这些有害的不被期望和不可预料事件的后果，我们采取的机制包括紧急预防和保险，它们都是我们谨慎对待未来的例子。

紧急预防和灾难预防

在我们家最近的个人环境，我们经常采取措施来预防灾难。我们安装烟雾监视器来提供火灾的早期预警，喷洒灭火系统来熄灭初始的火焰。我们也许有一个家庭内的安全报警系统，如果着火信号或未经许可的侵扰启动系统，系统就会自动呼叫消防部门或警方。我们也可能会保持电池、瓶装水和食物的储备以防断电期延长，甚至我们可能还有一个小型独立发电机能够紧急供电。要不然，如果我们不选择上面的预防措施，无论将来发生什么事情，我们只有碰运气了。

类似地，也有一些通过政府部门决定才能集体实施的措施。分区法律(zoning laws)能够限制在泛滥平原进行住宅建设，建筑法规对地震多发区的建筑材料及设计会有更高的标准要求。如果发生由地震、洪水、飓风引起的混乱，市政府需要指定和提供紧急避难所。国家可能会建立国家石油储备用于在外部资源可能中断的紧急情况下提供燃料。

对于“可能破坏诸如配电高压输电网这类非常复杂和高度分布的系统”的不可预料事件，需要进行特殊考虑。功能与设计的冗余(redundancy)是一种常用的方法。目前设计中的冗余对于在复杂系统中如何防止彻底失败具有决定性作用。在某些方面，冗余是“多米诺骨牌效应”(domino effect)的对立面，多米诺骨牌效应中，一长排中一个多米诺骨牌倒下，接着下一个会倾倒，再下一个。这种连

锁反应只有当所有多米诺骨牌倒下时才结束。在拥有足够冗余的系统中，一个位置的故障能够被隔离和回避，系统其他部分担当受损元件的角色，由此防止系统运行停止。

在旅行中，我们都曾遇到过道路或桥梁不通的情况，被迫走便道绕过这一问题，最终到达目的地。设计良好的交通网络通常具有充分的冗余以避免产生瓶颈，更不用说全部关闭了。设计良好的电力分配系统有充足的冗余来适应一个变压器的偶然失灵、一条地下电缆不经意的切断或者一个发电厂的紧急关闭。但是，偶尔所有多米诺骨牌倒下了，系统的不充分性在导致普遍断电的一连串故障中就显得非常突出了。

今天的城市通常都需要公共服务系统——煤气、水、电——在设计时具有很大程度的冗余，以便围绕受损区域部分的绕路很快起作用。事实上，电力分配系统中许多临时绕路相对是不需要人为干预的，当一个故障破坏电力分配系统时，它通过负载传感器和自激式交换装置发挥作用。这些预先的设计似乎是效果显著而明智的，但并不总是这样。1906 年旧金山的地震，除了震动引起建筑物的直接破坏之外，火灾是辅助破坏的一个重要原因。由于城市供水系统遭到地震的破坏，这种破坏变得更糟。这种情况导致灭火所需要的水量不足，原因在于没办法切断因总水管破裂而导致水流失的地区水的供应，这些水本来有助于扑灭其他地方的火灾。容易想象，今天这样的问题更加严重，又出现了不能切断向天然气管道发生破裂地区的天然气供应，在这些地区天然气作为燃料为火灾作了准备。

自然系统冗余的价值，特别是生物多样性，也逐渐受到更多的重视。生态系统中这种相互纠缠的特征有时只有在我们开始破坏它时

才变得很明显。在多样的自然系统中，对于碳和水来说，也许有多条路线沿着它们各自的循环通过这一系统。然而，当使用陆地的变化被强加于某一个地区时，这个地区的水文学模式被打破了，栖息地发生了变化或者是遭到了破坏，绕路的选择就更加小了。历史上伴随着土地开发的最初变化之一是湿地的排水，湿地又被人轻蔑地称为沼泽地。但是湿地作为候鸟迁徙栖息地的重要性和作为天然水净化和过滤系统的价值仅仅是很晚才被认识到。当我们不经意地破坏一个复杂生态系统的联系时，系统将面临着一个功能的退化。当不能进一步获得绕路时，此生态系统很可能最终崩溃。

但是科学家与工程师是如何识别复杂系统的微弱联系的，在我们生活的现实世界中他们又是怎样识别带来灾难的潜在瓶颈和多米诺骨牌呢？一种方法是采取“如果怎样”数值实验来努力理解这样一个事件的后果。“如果怎样”实验和复杂现象的大型计算机模型（见第七、第八章）的类别相一致，在复杂现象中系统的组成部分之间有许多的联系和反馈。系统的某个组成部分的随机干扰被引入，当它们传播到整个系统时，就会观察到结果。例如，电力公司玩“灾难性游戏”，模仿发电能力的意外受损或者传输网络分支的失灵。通过向不同位置的系统或者是包含不同组成过程的系统引入问题，实验者或者分析家有时会发现系统的弱点。计算机模拟能够检查故障发生的可能性，并确定每一种故障发生的概率，其中的每一种故障都会引起不同的后果。

试图预测问题并做出调整来消除这些问题取得巨大成功的故事之一是“Y2K”传奇。随着时间从1999年向2000年迈进，令人担心的计算机小故障引起了整个世界的混乱。在计算机使用的早期数据存储空间非常宝贵时，许多计算机程序在输入年份时截去前两位使用其

缩写形式。这样 1967 年变成 67，1994 年变成 94。但到了 20 世纪 90 年代世纪之交时，也就是 1999 年向 2000 年转型的时刻，其输入的模棱两可性是很明显的：02 的意思是 1902 年还是 2002 年呢？由日历决定的系统和程序——包括行程安排、工资、商品期货、退休金等无限长的列表，通常需要被修改以避免世纪末的混乱。由于问题本身出现的确切时刻总是知道的，因此对不确定性的测量被排除在方程之外。为了修改计算机使用早期采用的这种缩写，全世界耗费了许多人多年的心血和大量的金钱。日历翻转得准确无误，肯定是原先在补救方面付出大量努力的结果。然而人们也能听到如下观点：在日历翻转中之所以没有发生重大问题，其原因是对这一问题及其可能后果的描绘比实际情况严重得多。

虽然定量描述各种模型的组成部分的关系和相互作用仍然是一个挑战，但是对重大的不可预料事件的政治结构反应和社会结构反应能够通过计算机模型来研究。然而，战略和军事计划者建立和运行“战争游戏”模型，来检验社会结构、政治结构和军事结构中的各种概念化对包括随机事件在内的各种破坏的反应。在 1991 年的海湾战争计划中，利用计算机执行了许多模拟，来检验对可能发生的各种形式的反应。如果伊朗站在伊拉克一边参战会怎么样呢？如果利比亚发射一枚导弹攻击以色列将会怎样？一个持续的沙尘暴会如何影响美国坦克和装甲兵员运输车的机动性呢？浓厚的云层笼罩于地面部队的上空达一个星期对近距离空中支援的后果会是什么呢？在地面战争的摩擦中，如果伤亡人数增加，在美国，会对公众支持战争的努力产生什么样的影响呢？

当然，在战争游戏中，人们在对其进行探索之前，首先必须想象可能的场景。虽然第二次世界大战中日本敢死队飞行员的经历肯定

预示着这种可能性，但是，有关2001年9月11日的场景，满载乘客和燃油的客机故意与摩天大楼相撞，显然没有被认真考虑，甚至没有被想象。

年轻人花费无数时光来玩的许多计算机游戏之所以让他们感兴趣，原因在于他们每次玩时都面临不同的情况。尽管基本程序仍然相同，但每次计算机都改变了场所。同样，玩桥牌者之所以能够玩出无穷的技法，是因为每次都面临新的一手牌，计算机发牌能够随机生成新的组合。同理，现在计算机可以展示下棋时形成的很强技能——在对手(或计算机对手)每一步棋之后调整战略。较好的游戏者是，那些从上述多种经验吸取教训的人和开发出“对宽泛的环境成功作出反应”的游戏模式的人。

“如果怎样”实验同样也能对自然系统实施，例如，我们想要了解地球的气候系统是如何对大的火山喷发作出反应的。在历史记录和地质记录中有充分证据证明大的火山喷发对气候会有显著影响。1991年菲律宾皮纳图博火山喷发之后，在接下来的两年中，由于烟和灰阻挡了部分日光到达地球表面，全球平均温度下降了0.3℃(大约0.6℉)。

1815年，印度尼西亚的坦博拉火山发生剧烈的喷发，巨大的爆炸声在大约1200公里(750英里)外的雅加达都可以听到。坦博拉爆炸式地向大气投射了140多立方公里(35立方英里)的灰和浮尘，比176年以后的皮纳图博的爆发多10多倍。坦博拉火山爆发的下一年，即1816年，全世界范围被认为是“没有夏季的一年”。整个北半球温度降低，新英格兰州的一些地方下雪、湖水冻结，北美和欧洲的庄稼普遍歉收。英国诗人拜伦(Byron)勋爵，在瑞士度过了一个痛苦的暑假，写了《黑暗》这首诗，其中这样描述了坦博拉喷发带

给当地的影响：

明艳阳光为之减色，
星星在黑暗的太空漫游，
没有光亮、没有道路，冰封的地球
盲目地旋转，在无月光的空气中变暗；
清晨来了又去，去了再来，不可终日……

在6世纪，也发生了一个巨大的火山喷发，[1]欧洲和亚洲的历史记录对此有广泛记载。发生在公元536年的这次喷发的影响，由一位同时代作家[2]描述如下：

阳光变暗了，它的黑暗持续了18个月。每天它照射约4小时，而且这光仅仅是微弱的阴影。每个人都声称阳光将永远不能恢复它完全的亮度。水果熟不了，酒尝起来像酸葡萄。

在一个探索火山喷发对气候影响的"如果怎样"实验中，实验的变量包括喷入大气的火山灰和烟雾的体积、气体和烟雾中硫的含量，以及喷射的爆炸性，这种爆炸性将影响尘土达到大气的高度。其他重要的因素是喷发的地点、当时的季节和喷发的位置。由于温带的风循环模式与两极或赤道的不同，因此温带的喷发与两极或赤道的喷发相比，具有显著不同的效果，其冬季的喷发，将黑色的灰扩散到雪覆盖的白色领土，这将改变吸收阳光和反射阳光的平衡。系统的探索也许包括对地球上所有著名的多火山区的喷发、稳定进行的火山喷发对比单一的爆发事件、从火山释放出的气体的化学性质，还有众多

的其他因素进行实验。考察的场景的数目越大，我们对可能的气候结果了解到的也就越多，我们对不同地点、不同时间发生的事情进行预测就更加有信心。

保险

保险是一个很普通的概念，几乎对于想防止不受可能发生的灾难影响的每个人都很熟悉：汽车保险、房主保险、健康保险、人寿保险。我们向保险公司付年度的款项，反过来保险公司承诺，如果我们确实受到灾难打击，它们会付给我们一定数量的钱来补偿我们的损失，或者至少能够让我们平滑过渡到一个不同的未来。

某些种类的保险甚至是不自愿的。汽车责任保险就是一种不自愿的保险项目。在一些州，你可以选择不对自己的汽车损坏进行保险，但是你必须上一些保险来补偿在可能发生的交通事故中由于你的过错而遭受损坏或伤害的其他人。在拥有所谓"无过失"保险的其他州，不论过错怎么分摊，每一个车主不得不上一些保险。我称这种保险为不自愿保险，因为它经常是一个法律的规定或证明。为了获得驾驶汽车的执照，必须提供此证明。另一种不自愿的保险是房主保险，此保险至少是在这种情况下：一家出租机构掌握着一个抵押，由于此抵押家庭才被保证是有保障的。获得抵押的通常条件是借出方通过保险获得不受损失的保护。

对于给定类型的保险如何设立保险费(premium)？一个保险公司如何知道对于想参加本公司保险的顾客收取多少费用？保险价格建立的基础是对保险事件发生的可能性进行估计。这一可能性，最初并不是完全建立在过去经验的基础上。人寿保险是对于死亡所付的保险费，因此死亡率的统计为平均预期寿命提供了基线，同时也建

立了多少人遭遇过早死亡和多少人活过平均预期寿命的模式。对于历史上人口拥有充足的资料数据，这些统计对任何给定年龄(无论比平均年龄大或小)的死亡率给出了很好的估计。

保险公司和它们的保险精算师(在应用于保险业务的概率、统计、风险分析方面有专业技能的数学家)非常聪明，已经认识到通常的人口中由于职业不同而引起的死亡率的不同。例如，煤矿矿工和跳伞指导者也许比学校教师或会计付更多的人寿保险费。吸烟者常常比同龄的不吸烟者付的(人寿保险费)更多。同样地，地理上发生的预期寿命变化，是由于地区性环境健康因素；或者种族因素，例如，对患高血压的不同倾向性也许会导致在不同的种族人群中不同的死亡率统计。此外，由于保险精算师要对未来许多年的条件做预测，因此完成任务会更加困难。在死亡率方面，过去的经验是对未来的很好的估计吗？经济、饮食、大众健康方面的改善(或恶化)一定会影响未来的长寿吗？

相似的原则被应用于房主保险的定价。从基本的历史统计可获得今天的价格结构。在社区或县有多少家庭？一年中当地消防部门接到多少户的火灾呼救？这些简单的数字为每年家庭火灾发生的概率提供了一个基线。价格方面的第二个考虑也是基于你的财产是比平均结构的财产更多还是更少受损失。什么类型的结构最容易或最不容易发生火灾？你的房子离最近的消防站或消防龙头有多远？当然，保险费金额将取决于你房屋的尺寸、质量和家具，这些在估计房子替代价值方面是重要的因素。对于房主决定保险的其他危险，需要进行特殊的考虑。对于泛滥平原或易受地震灾害的地区，保险费也许会更高。

就像预期寿命随时间改变一样，影响财产保险费的许多因素同样

随时间改变。保险公司必须意识到随时间(变化的)趋势，并对其做出估计。通货膨胀率如何改变？变化的经济对社会财产价值具有什么样的影响？当“开放的农业土地成为人行街道和停车场”这一现象改变了水的过滤和排放模式时，城市蔓延对洪水模式和频率将会产生什么影响？通过复杂的数值模型对这些问题的评估与在第七、第八章讨论的情况非常类似。

但就像天气预报对未来短时间是非常有效的一样，为了预测保险花费所需要的人口统计学因素和经济因素，也是现在之后的一两年中最有效。当然，保险公司愿意花费更长时间窥视未来，但是处理随时间变化的数量的最安全策略，具有代表性的是对保险项目和保险费水平做年度调整，即一个短期的周期性“方向修正”。

在这些普通类型的危险保险(人寿保险、健康保险、财产保险，个人或公司购买它们作为对不确定未来的保护)之外，还有许多更专门化的保险产品：农业庄稼歉收保险、滑雪场降雪保险、商业中断保险、地震和洪水保险等。许多能够与天气和气候现象联系上。

就说在1997年春季北达科他州发生的一次洪水吧，此次洪水淹没了大福克斯城的大部分区域。当然市区商人会遭受由洪水破坏引起的财产损失，他们当中的一些人没有保险。但是他们的经济痛苦被以下事实放大和延伸了，他们在洪水退后的一段时期不能从事商业或提供服务。

或者考虑一下必须在春天种植庄稼，接着在生长的季节观察和等待谷物成熟的农民吧。由于降雨量不多，再加上降雨不及时，总体上庄稼产量可能会大幅度减少或受损。相似的，冬季娱乐区域的经营者依赖充足而及时的降雪。造雪机在没有雪时非常紧要，但是却不能替代丰富的自然降雪。结果是，农民和滑雪场经营者能够通过

专门化的保险保护他们不受天气不确定性和反复无常的危害。[3]

意外

自然界具有许多不可预测的活动让我们为之惊讶——火山喷发、地震、突然爆发的洪水。每年大约有4000次造成破坏的大地震在地球某些地方发生。尽管我们已经了解了相当多的关于地震发生的地理知识，但在一个给定地区什么时间可能发生地震？对此问题的估计，我们仍然没有取得任何真正的进步。

地震在地球表面不是随机发生的，但它们的发生也不是平均分布在地球上。一个世纪以来对“地震发生在何处”的观察显示，它们发生在特定的地理区域。居住在加利福尼亚的美国人都很熟悉声名狼藉的圣安德列亚斯断层，土耳其的市民已经通过灾难了解到贯穿该国北部的安纳托利亚断层。然而，沿着这些断层线，查明下一次灾难性事件发生在何时何地仍然是地震学家无法做到的事。对于地震事件发生时间的细致分析显示事件之间是没有规律性的。事实上，这些分析不能否认以下假说，地震之间（不论是地区性的还是全球的）的时间间隔完全是随机的。一次地震之后的下一次地震不定什么时候到来。

自然，地震学家已经了解了关于巨大地震之间的时间间隔的一些事情。沿着洛杉矶东部圣安德列亚斯断层，在加州南部，地质学家已经确定大地震破坏陆地表面是在1857年（在历史记录中作了很好文字说明的事件）、1745年、1470年、1245年、1190年、965年、860年、665年和545年。从中我们能够知道，在这9次沿着圣安德列亚斯断层发生的大地震的时间间隔平均大约为170年，但是最短的是55年，最长可达275年。这虽然能为“几十年之后，至少在加州南部，

也许会发生一次‘大的地震’”的忧虑提供一些试验性的支持，但对未来它不会是一个确切可靠的指导。但是何时何地发生大地震的不确定性或者就此而言发生任何有破坏性的地震的不确定性允许人民在地震活跃的地区建造房屋和生活。

对于地震造成的破坏用什么方法可以保护？由于地震何时发生非常的不确定，因此规定要交纳相应的保险费。在加州，地震保险仅由加利福尼亚地震局(CEA)提供，该单位是私人资助却实行公共管理的实体(其座右铭为“我们不能预测未来——我们只能保护它”)。CEA的年度保险费随地区而变化，但是通常在大厦成本(不是地皮)的0.1%到0.6%的范围内变动，同时要交纳保险赔偿金的一个非常大的可扣部分，即15%的保险赔偿金。价值为300 000美元的房屋，在交保险费之前，房主得先付45 000美元的维修费和重建费。显然，这里的保险只是保护不受大灾难的破坏，但是它仍然为房主留下了很多的风险和责任。

加利福尼亚也有洪水保险。初看起来，它与地震保险在定价上相似，每年缴的保险费占房屋价值的一小部分。同样对于上面提到的价值300 000美元的房屋，地震保险和洪水保险每年都要交约600美元的保险费。但是在扣除保险赔偿金方面有很大差别。对于洪水保险，房主在买保险之前只需支付500美元的洪水损失费，然而地震保险却需要支付45 000美元的可扣除赔偿金。对于同样的保险费，房主买的针对洪水的保护比针对地震的保护要多得多。对于地震保险有这么高的可扣除赔偿金，说明地质学家和保险商对于地震的了解要比对洪水的了解少得多；科学不确定性也反映在下面的规定中：对于地震保险房主要承担更高比例的风险。

CEA能够提供更小的可扣除赔偿金的政策吗？原则上是可以

的，但是价格当然更高了。如果可扣除赔偿金减少，保险费将以10的系数递增，那么一个人扣除4500美元每年就需要付6000美元的保险费。经过50年，人们付给保险公司的钱就相当于房屋的价值，而且在那种保险费水平上，这种政策也许称为长期自筹资金重建计划比称为保险保护更合适。

保险业处于风险评估的最前面，要尽力对未来趋势进行预测。这一行业减轻自然事件和人为事件对它的不利影响，同时以对损失和破坏进行合同补偿的方式，帮助顾客面对不确定未来。保险公司必须是认识人口统计状况变化、生活方式变化、建设标准变化和经济变化的第一批行业之一。在这一基础上，它们必须对它们的政策结构、花费水平和保险费水平做周期性调整。由于保险公司是第一批为变化的结果提供赔偿的行业，因此与其他行业相比，调整的余地更小一些。这对理解变化的含义来讲，也许能够提供一个更从容的“等待和观望”的方法。对变化采取思想上的抵制对于任何行业来说都是一种危险的态度，尤其是对那些通过保险承担责任的行业来说就更是如此。

气候意外

在气候变化的地方，我们了解到一个地区的气候变化有时是突然的、不可预料的。根据气候的定义，可以知道气候是温度、降雨、云量、风速和其他因素逐日变化的长期结果。因此，就突然的气候变化来说，我们更愿意大的变化发生在一二十年甚至一个世纪之后。在认识和理解突然性气候变化的原因方面，人们已给予了许多的关注，这些关注主要集中在海洋以及全球范围内洋流移动热量的方式等方面。绝大多数地理课程（无论是小学的还是大学的）都指出，由于

西欧和斯堪的纳维亚与加拿大的哈得孙湾大约在同一纬度，并且相当靠北，因此这两个地方的气候比人们预想的要温和很多。这样的情况之所以发生，是因为洋流、湾流将热带纬度加热的水输运到遥远的北方的大西洋，由此温暖了冰岛和临近的欧洲国家。

地质证据已经证实过去有一些时期，洋流似乎中断，全球热传送变慢或突然停止。对这一现象的理解，可形象化地通过百货商店的电动扶梯作初步而试探性的解释。电动扶梯的移动楼梯将购物者送到上面的楼层，但接着楼梯消失，返回到下面楼层。这些楼梯由看得见的部分和沿相反方向运行到我们下边看不见的另一部分组成一个大的环路。大西洋里的洋流也是如此。湾流沿着海的表面向北移动，但是也有一个相反的洋流沿着洋底向南移动。向北的表面洋流是从南部携带热量到达高纬度的暖流，向南的洋流是携带冰冷稠密的北极水从远道沿着洋底到达南极洲的寒流。无论是在百货商店的电梯里还是在海洋里，如果环路的一部分遭遇障碍，整个环便会断掉。

因此，是什么让洋流停下来的呢？首先我们必须知道什么是维持洋流循环的重要因素。驱动洋流的装置的一个必要组成部分是，从遥远的北方向南运动的冰冷稠密的底部水的形成。两个因素影响这种水的形成：温度和盐度(salinity)。温水不如冷水稠密，因而更有浮力；通常，它在海洋表面附近形成，吸收太阳的热量，在海洋表面行进。然而，冷水就更稠密一些，通常沿洋底行进。

盐度在水的密度方面同样扮演了一个引人注目的角色。海水盐度越高，密度就越大。相反，盐度越低，就越有浮力。因此，从江河流入海洋的淡水趋于呆在表面附近，直到它与含盐正常的海水很好地混合。然而，每年在北冰洋形成的冬季海冰把盐分排除在外，这些盐作为浓浓的盐水流沉入海底。在遥远的北大西洋，来自北冰洋

盆底排入到大西洋的冰冷盐水促进了稠密海底水的形成。这条路是北冰洋仅有的重要出口。接着稠密的底部水向南推进，由此完成了海洋传送带的循环。温度和盐度对洋流的综合影响导致了“热盐环流”（thermohaline circulation）这一技术术语的出现。

建立和维持洋流循环的关键链接是北大西洋“底部水”的形成过程。洋流是如何突然停止的？如果表面的水密度减小，它将“漂浮”而不是沉入到底部。这可以通过北冰洋和周围地区逐渐变暖的方式发生。更少被冰覆盖的海水，会直接受到阳光的温暖。同时，当北冰洋的冰冻层融化，通过向北流动的大江河将越来越多的淡水输送到北冰洋时，它的盐度会降低。这些向北流动的河流，包括灌溉亚洲大片地区的鄂毕河、勒拿河、叶尼塞河和通过加拿大流入北冰洋的马更些河。随着年度海冰面积的减少，浓盐水流就较少形成。这一过程中的某些时候，水将变得太有浮力而不能下沉，底部的水就会失去补充。由于传送在重要的转折点停止，由湾流向北的热传输将停止。

极具讽刺意义的是，北冰洋的变暖也许会导致海洋热传送带的中断，向西欧释放剧烈的寒气。预兆的信号已经出现：2000 年夏天，一艘加拿大破冰船目睹了北极点大片开放的水域，并未受阻碍地通过了西北的水路而没有遇见冰。2001 年夏天，又看到了通过这一遥远北方的另一段水路，一艘拖捞船从格陵兰岛到阿拉斯加进行了破纪录的航行，全程仅用了 17 天，船体并没有变得冰硬。[4]对北冰洋浮冰（ice pack）的水下测量显示 20 世纪后半世纪以来大约变薄了 40%，卫星图片显示同期内被冰覆盖的北冰洋面积已经减少。北极变暖的趋势是否会打断当前海洋循环的模式或什么时候将要打断，还不能准确地预测。对冰岛和欧洲之间某一深处洋流的海底测量已经显示 2001

年与 40 年前相比，洋流速度已减慢了一半。

以前曾有过一次，大约 11 000 年前，北大西洋的热盐环流发生了实质的变化。这一事件发生的原因可能是，当北美冰层向北消退到达一个地方时，冰冷的冰河融化水通过劳伦斯河谷到达海洋，而不是经由密西西比河向南流入墨西哥湾，从而将大量冰冷的淡水输入到了北大西洋。湾流的冷却和变淡相当快，发生仅需要几十年时间，持续却达 1000 年之久。西欧的温度下降了大约 5 ℃(9 ℉)。

这有点类似于在黑暗的山谷中行走，却没有意识到正在接近悬崖的边沿。每一步看起来都是正常的，直到从悬崖坠落。相似的，当你在爆玉米花时，热量慢慢地让玉米粒升温直到它们到达临界温度，在临界温度时突然会发生许多玉米粒爆开的行为。或者牛奶在瓶中未被注意地冷却直到它冰冻并扩张到顶起瓶盖或者撑破瓶子。在这后两个例子中的每一个，通过对系统的充分观察，我们可能认识到我们正在接近一个巨大变化的临界点。但是对于洋流行为，我们还没有充分的知识来了解我们离海洋热盐环流的悬崖距离还有多远。然而，意识到临近处有悬崖就非常有用了。

自然界不可预知的结果经常是作为巨大的意外出现。当自然系统突然朝一个全新而且明显稳定的行为发展或者发生剧烈运动甚或完全停止时，它们可能会以它们常见的行为方式，在设定的边界范围内慢慢变化，让系统看起来有些恢复。这种动力学系统从一个状态到另一状态的突然转变行为，已通过一个被称为复杂系统理论(complex systems theory)的数学分支进行了广泛的研究。

“保险”概念与上述的形势相关吗？显然这里的保险必须在国家或全球范围进行考虑。此外，为“保险”付费多少，也就是，设定保险费，并不是一项简单的工作。我们能够想象海洋循环的一个变

化所带来的全部后果吗？我们能够确认受影响最大的地区和国家吗？“保险费”应当在所有国家统一评估吗？建立反映哪一个国家受影响最严重的保险费应不应该做调整？应不应该有一个保险费的变化来反映哪个国家引起了该问题的不均衡？这些问题似乎不可能客观地回答，许多人试图与这些问题搏斗，由于受挫都绝望地放弃了。

然而，当我们面临气候改变以及它的可能后果这样大的话题时，有一个比拟可能会带给我们适度的安慰。世界上所有国家都要决定花费多少国家财富用于国防。简单意义上，军队组成国家的一个保险契约。原则上，它们保护国家，抵制外部侵略(和不时的内部叛乱)。当我们付房主保险费时，实际上我们希望钱被浪费掉，因为我们从不希望不可预料的事件发生。对于军队，同样我们宁愿从不调动它们采取行动，但我们常常得承认，至少在某种程度上它们是必要的。

为了保护国家不受外来威胁，需要多少保险费才是适当的？也就是说，合理确定国防开支是多少。对大多数人来讲，估计上述问题的答案是不可能的。不过，世界上大多数国家在它们的年度预算里确实有军队组成部分。除了极少数的例外，世界上的国家，无论大小穷富，没有不对这一保险进行投资的。任何国家选择花费的数量取决于许多因素，其中一些只是准理性的(quasi-rational)，与可能对手的力量的定量评估、长期建立的腐化和团体福利模式，以及诸如国家形象和自豪感这样一些明显的感情问题等相关。然而，最终我敢肯定，世界上的国家对国防进行投资的根本原因仅仅在于它们的国民相信这对国家安全是必要的。个人不可能很轻易地清楚阐明国家所需特定规模军队的详细原因，也不能确切知道军队应该拥有什么武器。那些细节是职业政治家、文职官员、将军和司令们的事情。但

是基本的判断已经做出：我们需要军队来保护。在纽约遭受恐怖袭击后，美国军队被召集在阿富汗采取行动，绝大多数美国人不会抱怨在军事准备上的开支。

同理，对于大多数普通人来说，要对“国家应付多少保险费来避免或准备全球气候的突变”这一问题进行理性评估也是不可能的。但是我确信，一旦国民理解气候确实在变化而且它会构成威胁，就会愿意花费一些国家财富制定补救和适应措施来减轻不确定未来的威胁。国民对此的研究不会比对国防的研究更详细。他们只是告诉政府和工业继续制定气候变化的保险政策，以某种比例形成补救和适应，全面认识到它的开支。

一盎司的预防……

在 1240 年，布雷克顿(Henry de Bracton)建议我们“一盎司的预防比得上一磅的治疗”。这一简单陈述里所包含的久经考验的智慧是足够清楚的：通常预防一个问题比后来处理它的后果要便宜得多。用在开发疫苗上的花费与治疗脊髓灰质炎、天花、黄热病、流行性腮腺炎、伤寒症、白喉的医疗和社会费用相比，要少很多。建设大楼时增设抗地震设施的花费比处理诸如 2001 年 1 月的大地震引起的灾难性城市崩溃的花费要便宜得多。2001 年 1 月的大地震毁掉了印度的普杰城，严重破坏了阿美达巴德。那次地震估计有 25 000 人到 100 000 人死亡，造成了 50 万灾民。在切尔诺贝利建立核反应堆时，制定更为安全的措施与处理 1986 年辐射发生的后果相比，效果要好得多，而且最终花费也会少很多。

逻辑无论如何清晰，也不会总是指导我们花费公共钱财或个人钱财的方式。医学研究更多地集中于治疗癌症，而不是预防癌症；医

学训练强调的是恢复健康而不是保持健康。甚至健康保险者，虽然用如下的标语掩饰自己，他们是健康维护组织(HMO)，然而历来却不情愿地为预防服务或诊断服务(如年度健康体检、节育、充分的产前护理、前列腺测试和乳腺X线照片)付费。现在，虽然HMO在认识这些措施花费的功效方面有了进步，但是他们宁愿为日益恶化的肺部疾患和哮喘付账，仍然对诸如空气质量等环境健康问题不够重视。上述情形在整个美国由于烟雾的原因已经达到流行的地步。

尽管这些大的健康保险机构看到(而且某种程度上也为其付费)环境恶化的健康后果，但它们仍然很少关注环境健康。其中一个重要原因是，它们的一些大客户来自促成环境恶化的工业。在我的家乡密歇根州，HMO在指责通用汽车或底特律爱迪生公司之前都要三思而后行，以免失去一个大雇主的健康护理业务。

从HMO的观点来看，也许不赞成环境和健康保护的决定有其自身的支持性逻辑。即使HMO对本地的顾客施加压力消除他们的行为，他们也感觉空气质量问题(和他们的保险花费)仍会继续。毕竟，密歇根的大多数空气污染来自伊利诺伊、威斯康星的燃煤发电厂和工厂以及其他远至西部的上风向的州。同样，密歇根州和俄亥俄州产生的许多空气污染会通过毗邻的州传到东部的宾夕法尼亚、纽约和新泽西。

现在，在美国西部的州的上空开始检测到亚洲的工业污染物，正一路行进穿越太平洋。我们发现这个问题确实是全球性问题。本地的HMO也许会简单认为这个问题太大而且远远超过他们掌控环境保护的能力，然而与此同时却失去在该斗争领域的客户。他们可能会争辩，全球空气质量是国家政府和国际关系应该解决的问题。

那些必须处理这些问题的政治家常常求助于科学不确定性并作为

维持现状的理由。他们很少独立从事任何科学研究，[5]这一点也不足为奇。像第二章提到的，许多人正是从那些破坏环境的工业获得科学观点，这些工业提供的信息是，这些科学证据实在是太不确定、太缺乏说服力，因而不能做影响未来的任何决定。当然，在这样的背景下，科学不确定性与社会保障制度所具有的让未来模糊的不确定性没有什么不同。如果我们的众议院议员和参议员在20世纪30年代采取了同样的态度，那么今天我们很可能不会有合适的社会保障制度了。这一争论一定与下边的疑问有许多相同的味道："当我们甚至不知道70年后有多少人活着的时候，如何制定那时的退休计划？那时他们的寿命会怎样？"尽管这些问题是合法的，但是如果社会保障制度的开发、实施或者无论何时提出这样的制度受到它们的阻碍，我们就不会有这样的计划了。恰恰因为国家在不确定性面前做决定是有准备的，我们现在才有了运行的制度。当然，今天的制度与20世纪30年代刚从国会出现的时候相比，肯定有着不同的结构。随着时间的变化，对其进行了几次方向修正以反映这种变化。

预防的原则：宁要安全而不要遗憾

本尼迪克在讨论臭氧问题的书中写道：[6]

> 当我们修建一座桥时，我们修建它要它承受的压力比它可能遇到的压力要大得多。然而，谈到保护全球的大气，赌注要高许多，这种观点（一些人的）看起来等同于要求把"桥将坍塌"作为加固它的理由是一件确定的事情。如果我们在制定保护臭氧层的措施方面犯了错误，那么我们应该意识到我们要对未来子孙负责、对预防方面所犯的错误负责。

当然你会承认采取预防措施的确是在讨论紧急预防和保险的那一章被描述过。本尼迪克对预防中所犯错误的评论作为预防原则的例子已经很出名。在气候变化的背景下，该观点于1992年在里约热内卢举行的联合国环境和发展会议(地球峰会)上被作了清晰的阐述。此次国际会议的一个成果是关于气候变化的纲领性条约，在条约中条款3规定，160个签约国家应该“采取预防措施来预测、防止或最小化引起气候变化的原因和减轻它的不利影响。在有严重的或不可逆转的破坏造成的威胁的地区，缺乏充分的科学确定性不应该被作为推迟该措施的原因……”

惠特曼作为新泽西州的州长(在被布什任命为环境保护局局长之前)进一步发展了预防的概念。在对美国国家科学院所作的讲演中，[7]她说：“我相信政策制定者需要对环境保护采取预防方法。对此我的意思是我们必须1)承认在管理自然资源中不确定性是固有的；2)认识到防止环境被破坏总是比以后修复它更容易；3)把提供证据的责任从提倡保护的那些人转移到提议有害行动的那些人身上。”她的第三个观点表明，说明一些行动的结果是有害的重担落到环境学专家肩上的时间太长了；现在她感到这一重担必须转由工业家来承担，以便在开展不会导致危害的行动之前，提供一个很有说服力的理由。

此外，她认识到科学运作与政策制定相比，时间尺度有着根本的不同。1997年在一次有毒的微生物爆发之后，一些水域的捕鱼被禁止，以防止可能被污染的海洋食品进入人类食物链。这种毒性是如何发展的？海洋食品是否有可能被污染？科学家无法立刻回答这些问题，对公共健康的保护也不能无限期等待。在某些区域禁止捕鱼的决定是依据手头最好的信息做出的。“缺乏确定性不是不做事情

的借口”，惠特曼州长说，后来又补充道，“如果我们想要好的科学，我们就不能冒犯它。”

后来的科学研究显示这种有毒微生物并没有感染海洋生物，但是州长为捕鱼禁令做辩解，她说，“我认为站在公众健康的一边犯错误是谨慎的，”由此清晰地表明了预防的原则。她承认这一决定已经造成了经济的下滑，但是她指出，如果污染和疾病发生了，对海洋食品和旅游业造成的有害的经济影响甚至会更严重。

科学不可能产生完全的确定性，肯定不可能完全按日程表进行。就像科学作家阿佩尔(David Appell)指出的：[8]“……现在，世界面临着极其复杂和严重的问题——全球变暖、生物多样性的丧失、环境中的毒素——这些问题科学还不能全部回答，传统的风险评估和管理不可能胜任这项工作。的确，就算知道这些问题的范围，它们也没有办法(回答)。”1999年，这一观点由气候科学家最早的专业团体——美国地球物理学联合会(AGU)在一份关于气候变化和温室气体的意见书[9]中进行了阐述：“根据地球气候系统的复杂性，对其描述和变化的预测永远不可能完全排除不确定性。”此意见书以下边的陈述结尾：“AGU认为，人类诱导的气候变化看似缓解，人类对气候变化也做出适应，但并没有证明科学不确定性的当前水平是不活动的。”

我们应该小心，不要让不完全的知识阻止预防措施的实施。我们不知道有关复杂系统的所有事情并不意味着我们对它一无所知。“直到我们有更多的答案才采取措施”的口号通常是那些坚持维持现状的人假装的一种微弱的借口。更多的研究并不是必然意味着有更少的不确定性——当我们对一个系统了解更多时，也许证明它要比我们想象得更为复杂，比我们最初想象的有着更大的不确定性。不论喜欢与否，我们必须在获得信息的基础之上及时做决定。尽管有时

真的憎恨浪费，但是的的确确什么都不做会更糟。在采取任何措施处理今天的问题之前，寻找有关未来的确定性是有害的规则。如果我们今天采取的措施后来证明是有缺点的，方向修正自然会被运用。

注　释：

1. 也许在巴布亚新几内亚东部的新不列颠岛会发现(Heming, R. F., *Bulletin of the Geological Society of America*, vol. 85, pp. 1253～1264, 1974)。

2. 也许是以弗所的约翰 (John)，但也可能是叙利亚人米迦勒 (Michael)(Rampino, M. R., Self, S., and Stothers, R. B., *Annual Reviews of Earth and Planetary Sciences*, vol. 16, pp. 73～99, 1988)。

3. 针对天气保险的一个有趣讨论，参见 James Surowiecki 的文章，刊载于 the *New Yorker*, 23 July 2001, p. 29。

4. *Chicago Tribune Magazine*, 30 September 2001.

5. 在美国第 107 届国会(2001～2003)上，只有两名成员拥有物理学高等学位。

6. *Ozone Diplomacy: New Directions in Safeguarding the Planet*, Harvard University Press, Boston, MA, 1991, p. 6.

7. Christine Todd Whitman, "Effective policy making: the role of good science", an address delivered on 13 October 2000 in Washington, DC.

8. *Scientific American*, January, 2001, p. 18.

9. *EOS Transactions of the American Geophysical Union*, vol. 80, n. 5, 1999.

第十二章　在一种不确定性的氛围里

我们如果以确定开始，就会以疑惑结束，如果以疑惑开始，并且耐心地解决疑惑，就会以确定结束。

——弗朗西斯·培根

我们走了很长一段路穿越不确定性花园，看到了许多花儿和一些杂草，它们美丽却有点杂乱。这个花园不是一个布置整齐的花园，呈几何形状展开，并趋于完美。花园里有许多隐蔽的幽深处，某些地方是让人吃惊的迷宫，每块土地都有我们从未见过的景色。当我们接近旅行的终点时，我们意识到不确定性就像花儿一样，在许多地方都能见到，多姿多彩。

在穿越不确定性花园的整个路途中，我们遇到了全球气候变化的许多方面：地球温度的测量、地区温度的变化趋势和全球的平均温度、洪水泛滥的概率、臭氧损耗、科学教育、工业宣传和困惑、媒体混乱、过去气候的重建、一周天气和百年气候的计算机模型、不确定

未来的保险。由于各方面不确定性与气候变化的不确定性具有相似的特性，因此在不确定性花园的每一个领域里，在探讨其他自然现象和人类活动的关系中必然会涉及它们与气候变化的联系。在最后的章节中，我会把引起气候变化的这些因素综合起来，提出它们自身的不确定性，说明在气候科学领域人们的努力奋斗和取得的成就，以及仍需攀越的高峰。

在20世纪的最后10年中，由于人们特别关注地球表面的温度，因此可能没有哪一个科学话题能够像全球气候变化这样引起人们广泛的讨论。关于这一重要且具有诱惑力的主题所进行的讨论是围绕着一系列问题展开的，其中每一个问题都可以单独提出和讨论，但是当把它们放在一起时，就包含了全球气候变暖整个的主题，并且内容相当全面。尤其是，这些问题及其答案说明了我们对全球气候系统理解的不确定性范围，另一方面也说明了不确定性所代表的挑战和机遇。每一个问题都包含着不同特征和大小及不同影响和强度的不确定性。下面让我提出这些问题并逐一讨论。

1. 在过去的一个世纪地球一直在变暖吗？
2. 如果是这样，变暖的原因是什么？
3. 这一变化带来的后果是什么？
4. 如何补救或者适应这一变化？

在过去的一个世纪地球一直在变暖吗？

为回答这一问题提供的基本证据我们称之为“仪器记录”。仪器记录主要包含大量由温度计测得的温度值。这些数据是由不均匀分布于大陆、海洋岛屿、固定或浮动的海洋浮标及海上船只上的气象台采集到的。在第五章中我们已经讨论过在进行这些个别测量并将

它们汇总计算出某一给定年份的全球平均温度时必须处理的许多问题。正是这些记录——在许多地方直接测量的温度值，表明了地球表面的大气温度在20世纪已经上升了约0.6℃(1.1℉)。这种上升是以其可能落在的温度范围出现的，这表明了温度数值的不确定性。其中上升范围在0.4～0.8℃之间有95%的概率。

实际上美国国家气候数据中心、美国国家航空航天局(NASA)和英国东英吉利亚大学气候研究小组，这3个不同研究组织都独立得出了同样的研究结果，因此我们对此结果就更加充满信心。对于如何调整以适应仪器使用的变化、如何平衡个别的温度值以获得全球的平均值、如何过滤出城市化的影响，每一个组织都确立了自己的质量控制措施。由于建筑物和铺就的道路具有吸收和保持热量的能力，因此城镇和城市会形成“热岛”，测得的温度值就会比此地如果没有城市化所测得的值偏高。在对气象站记录分析的早期，人们意识到城市化会使仪器记录产生向上的偏差。难道全球明显变暖不是仅仅因为许多气象站建在城市或城市附近吗?

幸运的是很容易排除这种疑惑。首先必须记住地球表面是由陆地和海洋组成的，它们大约分别占地球表面的30%和70%。很显然城市化不是海洋现象，然而，单就海洋表面温度而言，20世纪海洋上升的温度值与所估计的全球整体温度上升的幅度基本相同。第二，从陆地气象站得到的数据存在大量的冗余。每天都有5000多个气象站报告温度读数，有些气象站位于或靠近城市地区，有些位于农村地区。研究人员去除了来自城市气象站的数据，仅仅使用来自于农村气象站的数据进行分析。即使是只有大约2000个气象站是农村气象站，分析表明对于一个世纪的气候变化趋势，由仅仅根据农村的数据得到的评估结果与根据所有数据得到的评估结果没有根本性不同。

这就证实了对于城市记录的修正已有效地去除了“热岛”效应。结论是肯定的，全球变暖不是由于测量数据不准而产生的地理假象。

然而，在关于城市化效应的合理问题已经得到满意的处理很久以后，仍然不断出现怀疑的呼声，怀疑全球气候变暖的仪器记录是否完全排除了城市化效应的影响。这些呼声中很多来自于化石燃料工业，他们竭尽全力避免以下结论：煤和石油的燃烧很可能是导致升温的主要原因。他们的策略之一是通过坚称仪器记录受到了城市化效应的影响，否认全球正在变暖。

引起地表仪器记录不确定的另一原因是，集中在大气层 5 英里(8 千米)以下平均温度的测量是通过气象气球和位于地面 100 英里以上的轨道卫星两者共同确定的。过去 20 年以来的测量结果显示，这一位于地球表面以上的大气层区域变暖的程度比地球表面观察到的变暖程度明显要小。这种显著差异暗示地球表面的温度测量有错误。认真负责的科学家正确对待了这些差异，指出由于臭氧损耗，大气层位置越高其温度甚至越低。同时他们也指出大气中层和上层得到的观察结果并没有否定地球表面的观察结果。这些独立的观察结果，似乎都正确，它们指出有必要更好地理解是什么过程控制了大气层垂直温度的分布。

当然，地球表面正在变暖这一结论并不仅仅是根据仪器记录得出的。对于变暖，除了直接的仪器证据之外，科学家也观察到了由于变暖而导致的大量变化。给人印象最深的是，地球大部分地区的冰正在融化。在北极，海洋表面冻结成海冰，海冰的范围随季节而变化，但是在过去的半个世纪中，每年冰的平均面积缩小超过 10%。而且，在同样的时间内，海冰厚度的缩小就更加严重，大约为 40%。[1]如果气候继续变暖，按照现在的融化速度，很有可能北冰洋

将会在本世纪经历无冰的夏季。

南极也表现出巨大的变化。覆盖在大陆外围的漂流的巨大冰块以前所未有的速度脱落。自1995年以来，沿着南极洲半岛东边的拉森冰架、威德尔海中的菲尔希纳—龙尼冰架及大陆西边的罗斯冰架都减少了大量的冰。减少的尺寸令人惊愕：2002年拉森冰架减少的面积有美国罗得岛那么大，罗斯冰架分离出特拉华州那么大的面积。更大片的冰于1998年从龙尼冰架脱落，2000年从罗斯海脱落。罗斯海脱落的部分大约有瑞士面积的1/3！然而，并不是南极洲的所有地方都受到同样的影响。南极洲大陆东侧的一些小冰架却是稳定的，甚至稍有增长，沿着南极洲半岛的显著变暖在南极大陆其他一些区域并没有表现出来。

在气候更加温和的地区，高山冰川在其占据的山谷中回融到高海拔地区。实际上具有高山冰川的每一个地方都有明显的回融：欧洲的阿尔卑斯山、南美的安第斯山、新西兰、东非、喜马拉雅山脉、格陵兰、北美的落基山脉。有关以前许多冰川的大量历史观测资料和照片显示了在过去的一个世纪中冰川融化的程度。正如北极夏季的海冰一样，如果气候按照目前的变暖速度继续变暖，许多高山冰川将有可能在本世纪内消失。寓言中“乞力马扎罗山的雪”已经大大减少，也许它存在的时间不会超过20年了，[2]美国的国家冰川公园在本世纪的后半叶或许不再会有冰川。当变暖向山顶蔓延时，适应寒冷气候的生物被迫向上迁移，随之其他的植物和动物乘机占领扩展的气候较温暖的领地；传播疟疾的蚊子侵入到先前无疟疾的高山丘陵地带。[3]

每年对北半球的河流、湖泊冻结和融化的时间进行研究，结果揭示有朝着结冰时间推迟和融化时间提前的长期变化趋势，在上个世纪“冰季”时间减少了大约12天。[4]与之相对应，在过去的30年中阿拉

斯加的生长季节增加了 10 天。[5]自 1970 年以来，北美东部降雪区域的南部边界逐渐向北推移了大约 12 英里(20 公里)，每年温度在0 ℃以下的天数正在减少。 英国春季花儿开放和鸟儿产卵的时间提前了。 所有这些现象都说明多年来气候在持续不断变暖。

上个世纪海平面也上升了大约 4 英寸(10 厘米)。 海平面上升部分是因为陆地的冰融化成水最后汇集流入海洋，但是海平面的上升有 2/3 以上是因为海水的热膨胀造成的。 大部分物质都会受热膨胀，水也不例外。 研究人员费尽周折测量地球上各个海域的不同深度处水的温度，最深达 3000 米，得到一份海水温度数据表，从表中可以看出在过去的 50 年里海洋的热量显著增加了。[6]并不只是海洋表面变暖，而确实是整个全球海洋的水都在变暖。 海洋额外增加的热量与所观察到的被认为是海水热膨胀导致的海平面变化完全一致。

陆地表层下的岩石也有类似的证据证明气候变暖。 多年来我在地球物理学方面的研究包括在地壳钻孔以测量地球的温度。 在整个大陆钻了 700 多个孔，测量到的温度表明在过去的几个世纪里地球大部分地区几百米深处岩石已经在变暖。 研究发现 20 世纪岩石温度变暖与仪器记录显示的地球表面变暖完全一致。

每年行星的逐渐变暖，包括行星表面变暖、行星上水的变暖、行星的岩石地壳变暖，在很多方面都表现得很明显。 正是因为我们不是仅仅根据某种单一方式得到的观测报告，也不是仅仅根据一个研究组织得到的数据或仅仅根据有限的地理区域得到的观测报告，因此我们可以肯定地球变暖的确是事实。 人们无疑能够发现仪器设备、数据集或某人的统计分析存在问题。 但是当如此多的相互毫无关系的证据说明同一事实，那么客观的人们就认为事实是令人信服的，结论是不可推翻的。 因此，到了 21 世纪初关于地球是否正在变暖的争论

也就不复存在。现在主流大气科学家评估变暖这一事实的概率超过99%，地球变暖与一些对观察数据无法解释的曲解形成对比。换句话说，地球气候真的在变暖是一个“几乎确定”的事实。实际上，大部分的怀疑者现在承认20世纪地球已经变暖了。偶尔有人坚持认为此问题仍然是一个尚未解决的问题，这些人要么是来自扁平地球协会的有关气候学的人士，要么是来自以下这些人：他们所制定的经济日程或政治日程促使他们扭曲或者否认证据的充分性。

因此，当第一个问题得到明确回答以后，让我们进入到下一个问题。

变化的原因是什么？

事实上所有气候科学家都承认任何时候地球的气候是各种因素共同作用的结果。这些因素包括从太阳吸收能量的多少，地球如何反射或吸收能量，地球如何将能量释放到空间。通常气候系统运行的方式是，我们的行星从太阳吸收一定数量的辐射能量，吸收能量的多少取决于太阳的辐射量及行星与太阳的距离。行星距离太阳越近，吸收的能量越多；太阳系中离太阳越远的行星，吸收的能量就越少。由于辐射能随距离而衰减，通常太阳系中远离太阳的行星就会冷一些。

到达地球的辐射能大约有30%被地球表面尤其是白色的极地冰反射，剩下70%的太阳能被黑色的岩石、陆地的植被以及海洋中的水吸收，由此使地球变得暖和。但情况并不完全这样，因为如果是这样的话，由于太阳不断辐射、传送辐射能，行星不断吸收这些辐射能，那么我们的地球就会一天比一天变得暖和。地球是如何避免这种情况发生的呢？

众所周知，至少是在物理学者中大家都知道，所有物体都以相对于它们自身温度的某一速率发射能量。电气供暖装置辐射的热量在一定的距离内能够被感受到，整个下午沐浴在阳光中的建筑物墙面，到了晚上会辐射能量。与此类似，沐浴在阳光中的地球不会永远保留太阳能；它也会辐射它所吸收的太阳能。在地球演变的历史中，地球能调节它的表面温度使得地球正好辐射掉它从太阳接收的能量。这一表面温度代表着进来的太阳光和出去的“地球反射光”达到一种平衡。

但是气候系统比这种简单的平衡情况所描绘的要复杂得多。最重要的是，通过温室效应，地球的大气在地球—太阳的能量交换过程中起到了重要的屏障作用。大气温室效应的运行有点像捕捉动物的陷阱，进去容易出来难。大部分进入的太阳能很容易通过大气层，但是能量离开地球发散出去却很难。产生这种不对称性是由于进来和出去的能量所含的电磁光谱范围不同。来自太阳进入地球的辐射能主要位于光谱的“可见”波段(之所以这样称，是因为我们可以看见它)，然而从地球出去的辐射能主要是光谱中波长较长(且不可见)的红外部分。大气层对于绝大多数进来的能量来说是可以透射的，但是它却捕获一些向外辐射的能量。被捕获的能量会使地球温度比地球没有大气时的温度高。捕获机制是大气中的某些气体特别是水蒸气和二氧化碳能够吸收红外辐射。这些气体在大气层中只占微量，总量不超过大气的0.1%。但是绝不要低估这些微量物质的重要性，无论是分布于大气中的温室气体还是溶解于地下水的杀虫剂——它们都会产生极其深远的影响。

温室效应决不是其存在与否应予以讨论的某种理论概念，它是能够被观察到的真实物理效果。按照第六章提到的概率性语言，自然

温室的存在是一个“几乎确定”的事实。地球差不多从它诞生之日起就有温室大气层，而且我们应该庆幸地球有温室大气层。如果没有自然的温室，地球就会变得异常寒冷，海洋就会冻结，地球就不会是适于生存的地方。就这一最最基础的水平而言，太阳和大气的温室共同作用形成了地球长期的表面平均温度。

但是，太阳和大气的温室是如何引起20世纪观察到的地球短期的温度变化呢？问题的答案集中在，我们所观测的近期变暖过程，是从太阳获得的辐射能改变了还是温室的效应发生了改变。中心问题不在于太阳和温室效应是否影响气候——我们知道这两者都起着非常重要的作用。关键问题是在它们长期行为过程中，太阳和温室效应是否发生了改变，改变的程度有多大。近些年来，研究人员建立了重要的研究程序来确定太阳和温室效应的历史性变化，并将它们的变化与同一时间内地球表面温度如何变化的重建模型进行比较。这一研究的所有方面都是在不确定性的阴影下进行的，这种不确定性就像在第九章中讨论的一样，是重建过去的特征。

来自太阳的变化

几个世纪以前人们就已经知道，太阳表面经历着周期性变化；小黑斑出现了，并且数量增多，然后从太阳表面消失。这些黑子的出现和消失是太阳输出能量循环波动的证明。出现和消失的整个过程大约持续11年，然后不断重复这一过程。在最近的22个循环周期里，黑子的数量能够按天有规律地数清楚，大约400年前，黑子的变化很少有规律。在最近两次太阳黑子循环中，大约是在1979～2001年，地球轨道卫星一直忙于测量在每个循环中太阳输出能量变化了多少。测量的变化值作为“标度”用于根据过去太阳黑子数量的历史

记录估计太阳的辐射。

在系统计算太阳黑子方法之前，为了估计太阳辐射的波动，我们可以求助于其他的方法。由于某些化学同位素，如碳14和铍10，受太阳辐射作用，因此化石和岩石中这些元素的不同含量可用作太阳黑子数量或者通过卫星直接测量辐射能的替代品。同位素法能够让科学家估计几千年前的太阳变化。

温室的变化

正像对太阳变化的研究一样，过去40年里，世界许多地方用仪器直接测量大气中一些温室气体的浓度。这种测量是相当直接和常规的操作，既不复杂也没有争议。这些观测作为国际地球物理年计划的一部分，开始于20世纪50年代末。其观测结果显示，在整个观测期间，大气层中温室气体浓度稳定增加。在最近40年的观测中，二氧化碳的浓度比它在20世纪中期的水平增长了20%。再次运用概率性语言描述，温室气体浓度变化是一个“几乎确定”的事实。

然而温室气体浓度的记录，可以追溯到比最近半个世纪的直接测量要远得多的更早时期。由于冰中存有捕获的气泡（第九章提到的），这些气泡是大气中的“化石”，是气泡被捕获时的大气样本，因此科学家能够确定成千上万年前这些温室气体的浓度。气泡显示，在1750年，即大约工业革命开始时，大气中的二氧化碳水平大约为280 ppm，远远低于在20世纪末观测到的380 ppm。今天，在21世纪早期，二氧化碳每年以高于前工业时期0.5%的水平持续攀升。如果按照现在的速度继续增加，那么到2070年前后，大气中的二氧化碳将达到前工业时期的水平的两倍，也就是增加100%。

南极洲的厚冰保存了过去420 000年大气中二氧化碳的记录，这

样我们能够以真正长期的眼光关注最近的变化。南极洲的冰核显示当前二氧化碳的水平超过在长达约50万年的南极洲记录的任何时期的水平。而且，就像第九章提到的，这一长时期内的温度(就像冰中化学同位素所确定的)与二氧化碳的变化极度相关：二氧化碳浓度高，温度就高；二氧化碳浓度低，温度就低。如果这一相关性保持不变的话，以二氧化碳目前的水平和增长的速度，在本世纪后期高温是不可避免的。

二氧化碳不是浓度增加的唯一温室气体；甲烷的浓度也是前工业时期的水平的两倍；导致臭氧损耗的化学物质——氯氟碳化合物，在20世纪也发挥了重要的温室气体的作用。幸运的是，由于1987年《蒙特利尔议定书》提出逐渐停止使用氯氟碳化合物产品，大气中这些温室气体的浓度会缓慢降低。

大气微粒

除了来自太阳和温室气体的变化之外，其他因素在气候的变化中也起了作用。与大气中温室气体增加共同存在的另一重要因素是，被称为烟雾的微粒和小滴含量的增加。在工业化增长的过程中，化石燃料燃烧产生的烟雾和温室气体含量增多了。然而，烟雾却部分抵消了温室变暖效应，因为它们能在入射的太阳光到达地球表面之前将它反射回空间。与此类似，前面章节讨论的不时发生的火山喷发会释放灰尘和烟雾漂浮到大气中，这些烟雾和灰尘会在一两年内阻挡一些太阳光。

什么是改变气候的重要因素?

导致气候变化的因素有多种：变化的太阳、增加的温室气体、大

气中的烟雾、火山喷发物。这些因素中，我们能指出哪一种因素是导致最近气候变暖的原因？我们充分研究了每一种因素的历史，以便能够分析前工业时期及工业时期它们与地球温度之间的关系。在人类历史上前工业时期，温室处于稳定的自然水平，工业烟雾的排入量为零。因此我们必须注意用自然因素，如太阳和偶尔的火山喷发，来解释前工业时期气候的波动。小冰期的异常寒冷与长时间极少量的太阳黑子（1645～1715）的巧合，有力地说明了当时太阳的影响。1815年，坦博拉火山喷发之后“没有夏季的年份”证明了大气中火山喷发的烟雾和灰尘的那种虽然短暂但是又非常重要的影响。

然而，20世纪尤其是下半个世纪温度上升的趋势，不单单是因为自然因素的影响。事实上，仅仅使用自然因素如太阳变化和火山烟雾所得到的计算机模拟结果显示，由于火山喷发活动的好转，20世纪下半世纪温度实际上稍有降低。如果不考虑快速加剧的温室效应和伴随工业而来的烟雾因素，对20世纪温度的模拟会显著低于实际观测到的温度。只有加上工业烟雾和由于燃烧化石燃料形成的温室气体的增加这些因素，模拟仿真才能再现现实并与观测到的变暖相吻合。正是在气候模拟的基础上，2001年，IPCC的科学家得出结论：过去50年观测到的气候变暖主要是由于人类活动的影响。

这一结论具有多大的不确定性？IPCC评估上述结论是正确的概率在“非常有可能”范围内，即90%～99%的确定性。然而，当温度升高和温室气体浓度增加这些观测结果被认定是“几乎确定”的事实时，在复杂的计算机模拟中定量两者间的关系得出结论的确定性比孤立地考虑单一因素得出结论的确定性要低。仅仅在5年以前，IPCC的态度比较谨慎，他们认为证据之间的均衡仅仅说明了可以察觉的人类对气候的影响，真实的概率是66%～90%，它只是在“很有

可能”的范围内。经过其后几年的研究，IPCC 对于气候变暖的原因，把握度增加了，不确定性减少了。

这一变化带来的后果是什么？

预测气候变化对未来产生的影响，让我们先撇开重建过去本身的不确定性，进入到预测未来的不确定性任务当中。就像我们已经看到的，预测的不确定性有时是因为不能完全理解系统是如何工作的，有时是因为即使系统模型恰当，却不能预测对系统而言非常重要的因素是如何随着时间而发展的。第七章讲到美国社会保障制度的例子，就说明了由于不知道影响系统的重要因素是如何随时间变化而导致预测的不确定性；虽然相对而言控制系统的方程简单易懂，但是人口统计预测却需要随时修正。

然而，撇开不确定性的问题不说，有时人们首先关心的是另一个“谁来关心”的问题，而不是全球气候变暖后果评估的问题。如果变暖带来的结果都是有利的，那么就没有理由关心，也没有必要采取补救行动。因此，一小部分高声“反对”全球变暖的人提出“谁来关心”这样的问题，也就不足为怪了。由于目前对过去一个世纪中气候变暖的记录已经是“几乎确定”的事实，以及过去半个世纪气候变暖很可能是由人类活动而引起，因此，怀疑者退却到了第三道防线。他们告诉我们不必担心，全球变暖对我们有利。他们的观点是：天气暖和会延长各地的生长季节，大气中二氧化碳浓度增加会使生物长得更好更快，再加上温和的冬季，每个人都会很富裕。所有这些听起来很吸引人。然而，在这些表面现象之下，现实世界的复杂性大大减弱了对这种观点的支持。

植物生长加速本身就是喜忧参半的事情，因为生长对植物而言是

无可选择的；野草也会与作物生长并驾齐驱(或者超过？)，这就需要更多的人力和除草剂来控制野草生长。而且并不是所有的植物物种都受益，因此，占支配地位的物种控制领地就会取代其他物种。而且，植物生长加速需要的不仅仅是二氧化碳，还需要其他的营养物质，特别是氮化合物，当然还有水，才能保证产量增加。为了利用大气中增加的二氧化碳，就需要增加使用其他的肥料，随之而来就会出现江河湖泊水域的水质问题。既然大气变暖必然会在很多地方逐渐导致土壤干旱，那么二氧化碳的营养作用也会因土壤水分流失而大大削弱。植物生长的试验中，二氧化碳含量增多，其他因素控制不变，生物量确实增多，但是营养价值不变，这样营养就会由于生物量的增多，分配更加稀疏。

在阿拉斯加北部一处森林里，科学家寻找由于二氧化碳增多而树木加速生长的证据。他们发现的却是因为当地经历气候变暖而出现的森林衰落的景象。森林可以吸收大量的二氧化碳，但是由于邻近地区的高温却减弱了森林的吸收。此外，如果组成森林的主要树种衰退，其他的物种就会趁机侵入占领土地。由于各个物种都寻找机会抢占新的领地，森林将会上演抢座位游戏；然而，正如在抢座位游戏中一样，有些物种会发现根本抢占不到新的空间。因此从生态学的观点来看，在二氧化碳增加的条件下，一切生物原位不动，只是生长加快，这是很幼稚的想法。

后果评估的基础是对升温范围的估计。21 世纪研究升温的计算机模型得出的温度升高范围大约为 1.5～5.5 ℃(2.7～9.9 ℉)，这反映出一个相当大范围的不确定性。但是不确定性是关于什么的呢？基本层次上，这个不确定性范围的产生不是由于我们对全球气候系统的科学理解不确定，相当程度上是由于全球人口如何增长、地球居民

使用什么能源和多少能源的不确定。这些是有关未来社会和经济的不确定性，与第七章讨论的未来社会保障制度需要的不确定性具有相同的特征。

世界各地的大学、气候研究实验室都独立研制了关于全球气候变化的复杂的计算机模型。假如人口统计学和经济学在这些模型中是相同的，则由这些大量的气候模型得到的有关未来的气候预测结果应当是相似的。这么多研究组织在他们的模型中对于如何表达气候的物理原理和化学原理做出了独立判断，并对气候系统如何发展得出了极为相似的结论，那么我们应当对这些结论有一定的把握。有关未来气候的更大的不确定性不是源于气候科学而是源于社会科学和人类反复无常的行为。

不幸的是，全球升温带来的结果显然并不都是有益的，有些在前面作为升温的间接证据已经提到过。全球升温带来的最普遍的结果就是海平面上升，绝对没有人认为这是有益的。上一世纪海平面上升了约 4 英寸(10 厘米)，主要是由于气候变暖和随之而来的海水热膨胀造成的。根据不同的人口及其发展图景预测出的温度上升范围很容易被转化为 21 世纪海平面可能的变化范围：在现在的基础上再增加 8～28 英寸(20～70 厘米)。但是海平面上升的真正影响可以通过土地被淹没的数量体现出来。在一个稍微倾斜的海岸平原地区，如美国东部海岸和海湾海岸、亚马孙河、刚果、依洛瓦底河的入海口，海平面上升很小也会导致内陆海岸线很大的推移，淹没大片区域。在大多数关于 21 世纪的图景中，南佛罗里达的大多数地区都面临着被洪水淹没的危险，太平洋的很多居民岛屿也面临同样的命运。海平面上升意味着世界靠海的所有城市要完全改变：纽约、迈阿密、新奥尔良、阿姆斯特丹、哥本哈根、东京、布宜诺斯艾利斯等不计其

数的城市。修筑堤坝将极其昂贵，并且它们最终也会没有用。在像荷兰这样的小国之外修筑围墙是一回事，绕着整个的大陆建造和维护堤坝则完全是另一回事！

较低大气层变暖的另一结果是大气中水蒸气含量增加了。大气中水蒸气含量增加对气候系统意味着什么，对此引起的争议相当大。如果水蒸气变成更厚更多的云，云层就会将更多的入射太阳光反射到空间，这样可能会减慢或者稳定变暖过程。但是水蒸气也是一种温室气体，大气中水蒸气增多也可能会加剧由二氧化碳和其他温室气体引起的温度上升。大气中水蒸气的这种矛盾行为将如何结束是气候科学中较大的一种不确定性，但是由于它可能向两个方面倾斜，因此不确定性在概率范围内呈平方性降低。

另一方面，大气中水蒸气含量增多也增加了暴雨的可能性——24小时内降雨量超过2英寸(5厘米)都可以称为暴雨。对上个世纪国际天气文档详细调查[7]显示内陆地区经历暴雨有上升的趋势。2000年北达科他东北部地区12小时内降雨量达1英尺(30厘米)，新泽西中北部地区在24小时内降雨量达14英寸(35厘米)，大大超过了这些地区的历史记录。

估计出全球平均温度上升和全球大气中水蒸气含量增加的概率很高是一回事，预测出对平均值的偏离如何在全球范围分配却是另外一回事。任何平均值的增加都是由高于平均值的地区和低于平均值的地区共同构成的。结果，即使很有把握估计出全球范围的平均变化，但是预测地区范围的变化，不确定性却要高很多。不同的计算机模型，尽管都支持全球性变化趋势，但就地区变化趋势而言，结论却不一致。在北美五大湖区，有的计算机模型预测五大湖区水位会下降，但也有一些预测会上升。

在前面的章节中，我也讨论了气候系统的意外。当达到改变大西洋热盐环流的某一重要临界值时，气候就有可能突然发生变化，停止向北欧传输热量。或者当格陵兰岛大块的冰帽滑进海洋，海平面就会瞬间上升，就像把冰块扔进一杯水中一样。气候科学家仍然不知道如何确定这些事件的临界值，不确定性依旧不可避免。

如何补救或者适应这一变化？

现在我们进入了一个新的领域，一个并非只有科学性的地形。如何应对变化的气候？对这一问题的许多方面进行处理会迅速导致科学、经济、政治、宗教等多方面结合。这些相关领域的支持者，在未来的几十年内，是如何相遇的？他们是如何跨越这些“小山丘和山谷”似的障碍，又是如何在这个新领域彼此对抗的呢？关于这些问题的许多方面可以被描述而且将要描述。

当然，也有坚决的现状拥护者，如那些主张无论气候如何变化我们都不应当予以关注的企业家和政治家。这群全球变暖的反对者就该领域的问题四处散发小传单，规劝任何人甚至所有人最重要的事情就是“照常工作”，即保持现状。他们的观点大家现在也很熟悉：全球变暖只是一个理论；预测的结果非常不确定；科学家也不是意见一致的；数据会有错误；需要进一步研究。他们采取经典的恐吓策略，预言会发生可怕的经济后果，告诉人们如果采取措施对付气候变暖，他们将工作效率很低，并且在此过程中经济会遭受损失，人们会失业。正如我刚刚所讲的，有些人会说，气候变化是可以预料的事情，而不是不可控制的。在他们尽力创造的不确定性背后，人们很容易看出他们急迫关心的事情：他们认为为了应对变化的气候被提议的许多变化会给他们造成不利的经济后果。公众，特别是媒体，必

须认清这些观点的出发点明显是自私的，应当对它们适当权衡。

实际上有很多技术上可行的方法可以减慢大气中温室气体的形成。第一种方法是简单的能源节约。富兰克林(Benjamin Franklin)有关“节省的一分钱就是赚得的一分钱”的观点很贴切。节省1千瓦时的电就意味着不需要生产或购买1千瓦时的电能。加利福尼亚的居民采取他们短期内唯一可行的方法——能源节约，来缓解2001年的电能短缺的危机。节约措施的成功，尤其是其成效的立竿见影，让许多人大为惊讶。

稍微长远一点，可以通过很多途径来节约能源，如提高我们所用车辆和设备的效率。美国市场上畅销一种高效混合电能汽车，每加仑汽油能跑60英里(每升汽油25公里)，不需要给固定电池充电。高效的鼓风炉和空调、冰箱、洗衣机、干燥机、热水器正成功地推向市场，部分原因在于它们的整个生命周期内，能源成本的节约量增大。另外，高标准的隔热建筑对于节约能源来说很有潜力。这些措施不需要额外的研究和开发，现在的市场上已经有这些产品，马上可以用于减少温室气体排放到大气层。对股东来说，它们更有利可图，并且可以提供大量的就业机会。

另一类技术补救方法是开发不产生温室气体的新能源替代化石燃料。一些能源，像风力发电和水力发电，很久以前大家就知道并在一些地区应用。如今，风能的利用被看成一种技术复兴，利用它可产生大量电力。核能虽不污染大气，但是也有其他的环境障碍需要克服，这些障碍使得它替代化石燃料的过程大大减慢。2001年美国的“9·11”恐怖袭击事件之后，人们对现代社会在恐怖主义面前的脆弱性有了新的认识，这为核能的使用带来了必须要克服的进一步障碍：要安全可靠回收来自核能生成过程中产生的裂变废弃物，保护核

电站不受内部破坏和外部袭击。而在热岩石离地表较近的地区，地热能更具吸引力。在不远的将来直接用太阳能电池和氢燃料电池发电将大有前途。

缓和温室效应的第三种技术方法是回收温室气体，实际上就是在温室气体进入大气之前收集和贮存它们。将二氧化碳送回到地球表面(北海下用于开采石油和天然气的蓄水池)的试验项目已被证明是相当有前途的方法。另一种试验性的方法是把二氧化碳抽吸到海底，由海水吸收存储。然而有人批评海洋存储具有潜在的不稳定性，而且可能对海洋生态系统产生损害。至今，科学家认为进行光合作用的植物利用二氧化碳进行光合作用也是一种消除温室气体的方法，即将二氧化碳暂时贮存在生物中。虽然伊甸园的恢复并不像某些人所说是二氧化碳排放带来的益处，但是反对在发展中国家进行大范围森林采伐，在某些地区再造森林也是解决方案的一部分。

然而，变化是不可避免的，事实上变化正在进行之中。现在地球正对自工业革命开始以来发生的变化作出相应的反应。就像让巨型战舰下海，在下一个50年中，地球也会对我们所造成的变化反应逐渐变慢。无论是自然系统还是社会系统都存在着巨大的惯性。21世纪的前50年里，地球上的人口以及随之而来的需求会继续增长，甚至会超过预期。海洋吸收的巨大热量使得已经发生的变暖在未来的几十年里会继续影响全球气候。有关温室气体的《京都议定书》要求减少二氧化碳的排放，使之稍微低于1990年的水平，这也仅仅是减慢了温室气体潜能的增长，并没有减少温室气体量。已经存在于大气中的温室气体的存在时间，甚至不算任何新的增加，也应以数十年或数百年来计算。不管喜欢与否，地球上的人类必须对由于20世纪化石燃料使用和人口增长所引起的温室气体增加导致的变化做

实际预测。

因此，适应变化只是通向变化的未来的一种途径。对此能够做什么呢？让我指出确实需要注意的一些领域。面对海平面上升的问题，沿海陆地使用的传统方法将会遭到更大的困难，政府对于投资多少公共财产用在保护措施上面临着重大的抉择。洪水泛滥的低洼地区，大量的人被迫迁移，就像战争引起难民的迁移。公众健康活动也不得不进行调整以适应更大范围的热带疾病以及与热相关的健康问题。水将成为21世纪最重要的资源，对水的竞争将超过20世纪对能源的竞争。实际上，对淡水的需求可能最终导致太阳能在经济上最具竞争力，因为太阳能是让海水大规模脱盐淡化的主要能量资源。已经增长的淡水价格[8]将超过用来获取和分送淡水的增加的能源成本。面对必然的全球变暖引起的降雨模式和土壤湿度的改变，全球农业的地理也会相应改变，国家和国际水政策将走向前沿，成为政治和经济论坛争论的中心。

世界各国政府对这些问题会发挥什么作用呢？一是减缓导致气候进入新的未知领域的人类活动的巨大影响，二是对不可避免的变化使得地球人类做好准备。在各国范围内，有能够对温室气体排放切实施加影响的政府性行动。政府有权采取更广泛的税收政策，鼓励替代能源的研发，抑制对化石燃料的严重依赖。征收汽油税直接影响车辆燃料的效率，汽油的费用越高，人们就越希望有高效节能的车辆。因为该政策更加严重地影响了部分不太富裕的人，偶尔有人批评对汽油消费直接征税是一种社会倒退。然而，税收法的其他条款弱化了这样的批评。政府有权制定车辆燃油效率的标准，但是在美国出于某种目的不愿运用这种权力。2002年，美国参议院再次有机会强制建立车辆高效燃油效率标准，但是，屈从于生产汽车的那些州

的参议员的压力，他们拒绝这么做。

世界各国也有采取集体合作的机会。尽管就全球环境问题国际间的合作仍处于初级阶段，但已经有了一个比较突出的成功案例：《蒙特利尔议定书》，它是1987年磋商达成的国际性的协议，要求逐步停止使用氯氟碳化合物产品，停止破坏平流层的臭氧。协议签署仅仅15年之后，大气中氯氟碳化合物的水平已经稳定下来，不久将开始下降。尽管臭氧洞依然存在，科学家们预期随着氯氟碳化合物浓度下降，臭氧层将逐渐恢复。如果一切按协议进行，那么臭氧的分布在21世纪中期将恢复到氯氟碳化合物破坏前的状态。

由于氯氟碳化合物气体本身是温室气体，它们的消除在减少大气温室气体方面是前进了一小步。《京都议定书》是在1997年签署的关于开始减少二氧化碳排放的类似的国际公约，由于美国不愿参与而被迫回到缓慢而摇摆不定的初期阶段。由全球气候隐含的另一个真正的国际主题是全球人口问题。我们不应该忘记整个的能源消费是每个人消耗的能源乘以人口的数目。地区性和全球性的人口稳定，是减缓能源消耗增长的最有效的途径之一。它是在争论气候变化时从未谈论的话题。

危　机

上面符号是crisis的中文表示，它由两个词组成——“危险”(danger)和“机会”(opportunity)。由于人类活动导致的地球气候变化被视为一种危机(crisis)，在英语中是困境(predicament)的意思，汉语中却是风险和机会的意思。

如今人们广泛讨论并合理理解了全球变暖的“危险”——可是这个“危险”只是汉语“危机”一词的一半。关于“困境”一词的含

义，世界上很多地方趋于一致：人类正改变着他们生存的环境。然而，人们对全球变暖带来的“机会”——汉语“危机”的另一半意思，重视得很少。环境后果促使工业界检查自己的工作，发现能源、水及废物处理的花费可以节约很多。发展可持续能源就是机会，这有希望为开发可持续能源的人们带来丰厚的回报。与20世纪人们可负担起的汽车带来的改变类似，21世纪环境方面将会发生良好的交通模式的改变。一些有远见的工业已经认识到有利于环境的活动和产品为它们提供了强大的竞争优势。

我感到，在一些企业、政府以及更广泛的普通公众中，关于气候变化的更明智、更乐观的观点正在慢慢形成。这种观点识别出但并未被与气候变化相联系的不确定性阻碍。新兴的观点认为全球气候变化包含风险和机会，气候变化的缓解和补救必涉及技术创新和社会创新。自然，达成像美国布什政府关于能源政策、温室气体排放和应对气候变化的国际战略这样的共识，不时会遇到挫折。但是全球变暖的现实会无情地向前发展，最终将迫使一些决策者放弃他们的意识形态和政治惯例。

有时可以听到这样的说法，人要么是变化的作用者，要么是变化的受害者。尽管我们都是气候变暖的受害者，只有一些最严重的后果才会让人们想到一些补救，如海平面上升迫使人们迁徙。导致巨大变化的人中有许多是那些拒绝承认气候变化或拒绝对此负责的人。最终他们也会以某种特殊的方式成为气候变化的受害者。他们忙于保持现状，看不到在生活环境中进行的巨大的不经意的人类试验所带来的机会。他们保持现状的需求和对未来不确定性的恐惧让他们抱有偏见，转移了他们利用不确定性带给他们机遇的视线。他们看到玻璃杯是半空的，由于害怕玻璃杯会完全变空结果采取保守的态度。

尼赫鲁*(Jawaharlal Nehru)对政治上的怯懦所作的评论用在此处非常合适，他评论到，“过于谨慎的政策是最危险的。”

看到玻璃杯半满的人也将是变化的成功人士。他们会看到填满杯子的机会。不确定性会激励他们而不会吓倒他们，他们是通过创造未来而预测未来的人。

注 释：

1. Rothrock, D. A., Yu, Y., and Maykut, G. A., *Geophysical Research Letters* vol. 26, pp. 3469～3472, 1999.

2. *New York Times*, 19 February 2001.

3. Epstein, P. R., *Scientific American*, pp. 50～57, August 2000.

4. Magnuson, J. J. et al., *Science*, vol. 289, pp. 1743～1746, 2000.

5. Running, S. W. et al., *EOS Transactions American Geophysical Union*, vol. 80, n. 19, 1999.

6. Levitus, S. et al., *Science*, vol. 292, pp. 267～270, 2001.

7. Karl, T. R., Knight, R. W., and Plummer, N., Trends in high frequency climate variability in the twentieth century, *Nature* vol. 377, pp. 217～220, 1995.

8. 在地球的许多地方一升瓶装水的价格已经同一升汽油的价格相同了。

* 尼赫鲁(1889～1964)，印度开国总理，执政17年。——译者

译 后 记

近日，我将《不确定的科学与不确定的世界》的第二轮校对稿发往上海科技教育出版社之后，终于如释重负。正式拿到这本书的原著是2004年9月份，至今将近一年。回想起这段刻骨铭心的岁月，不由得让人感慨万千，好久了，自己竟然没逛过一次街，没添过一件新衣服，和许多老朋友未通过一次电话，累的时候经常是晚上没有力气洗脸倒头就睡。为了确保质量和进度，今年春节，我未能回家看父母。即使日夜追赶，临到开学，还是没有做完，这样原本想在2005年3月初就策划好方案、在4月18日全球激光束传递这样很有纪念意义的日子里开幕的物理文化节，只能一拖再拖，最终开幕的时间已经是5月27日。但不管怎么样，这本书的翻译成为我一年来生活的主要内容之一，我为此而喜悦、激动、困惑、痛苦、感叹，其中所经历的种种复杂感受只有认真从事过翻译的人才能体味。

有人说，翻译只要外语好就行，比写东西要容易得多，可能我以前也这么认为，但是真正翻译完毕，才知译事远非易事。以前我曾

和老师合作翻译过柯瓦雷的《伽利略研究》一书，这也是我有信心承担本书翻译的基础，但本次是独立翻译，在翻译过程中总是感觉困难重重。2004年9月份拿到书的原文，其后由于参加了一个不能缺席的培训班，真正动笔翻译的时间是在11月底。由于本书涉及的领域非常广泛，既有数学、物理、化学、天文、生物、地质、计算机等自然科学知识，也有经济、教育、媒体、保险、政治、环境等社会科学知识。书中还有大量的名人名言、谚语和广阔的背景知识。所有这些都为翻译增添了不小的难度。在这种情况下，我制定了周密的翻译计划，无论多么困难，都要把这件事情做好。对于书中涉及的对我国读者来说比较陌生的许多事件、人物和知识点，为了利于读者理解，我查阅了大量的相关书籍和网站，加了必要的译者注，核准了所有的人名、地名，并对照原文反复作了校正。

在本书的翻译过程中，得到了好多朋友的帮助，我首先得感谢清华大学人文学院刘兵教授、上海科技教育出版社的潘涛博士，为我提供了这项愉悦而有意义的工作。要感谢我的男友徐春波和妹妹李晓妮，是他们在学习、工作繁忙的情况下抽空为我查阅文献、输入全部译稿，并对全文的大多数章节进行了初步校对。还要感谢同事李西元、于芳芳和好友曹县芳，帮我校对了作者自我介绍、第四章、第五章、第六章和第十二章的译稿。另外，在翻译方面很有经验的于卉芹副教授、史焱副教授在翻译技巧方面为我作了指点，让我受益匪浅，对此，我感激万分。还有，装甲兵工程学院政教室副主任李清津先生，对书中所涉及的经济学术语的正确译法给予了指点；翻译室王惠仁教授对第十一章有关保险方面给予了详尽分析；外语室朱娜老师为我提供了原文引用的令我困惑的法文的英文解释，在此一并表示感谢。当然，我还应该感谢我的领导杨万利博士和刘家福博士，没

有他们的支持，我不可能去上海开会，也就不可能认识引荐我做翻译的潘涛博士。最后，我非常感谢上海科技教育出版社的编辑侯慧菊和傅勇，虽然我和他们未曾谋面，但本书的翻译和修改就是通过京沪之间的电子邮件、信件和电话这样的交流方式进行的，他们对本职工作的敬业精神和认真负责的态度让我非常感动，感谢他们为本书的出版所作的种种努力。

由于译者水平有限，加上时间仓促，译文肯定还有可商榷之处，期待读者批评指正。

李萍萍

2005年7月于北京